2013-09

Secure military communications on 3G, 4G and WiMAX

Schoinas, Panagiotis

Monterey, California: Naval Postgraduate School

NAVAL POSTGRADUATE SCHOOL

MONTEREY, CALIFORNIA

THESIS

SECURE MILITARY COMMUNICATIONS ON 3G, 4G AND WIMAX

by

Panagiotis Schoinas

September 2013

Thesis Advisor: Gurminder Singh
Co-Advisor: John H. Gibson

Approved for public release; distribution is unlimited

THIS PAGE INTENTIONALLY LEFT BLANK

REPORT DOCUMENTATION PAGE

Form Approved OMB No. 0704-0188

1. AGENCY USE ONLY *(Leave blank)*	**2. REPORT DATE** September 2013	**3. REPORT TYPE AND DATES COVERED** Master's Thesis

4. TITLE AND SUBTITLE
SECURE MILITARY COMMUNICATIONS ON 3G, 4G AND WIMAX

5. FUNDING NUMBERS

6. AUTHOR(S) Panagiotis Schoinas

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
Naval Postgraduate School
Monterey, CA 93943-5000

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING /MONITORING AGENCY NAME(S) AND ADDRESS(ES)
N/A

10. SPONSORING/MONITORING AGENCY REPORT NUMBER

11. SUPPLEMENTARY NOTES The views expressed in this thesis are those of the author and do not reflect the official policy or position of the Department of Defense or the U.S. Government. IRB Protocol number ____N/A____.

12a. DISTRIBUTION / AVAILABILITY STATEMENT
Approved for public release; distribution is unlimited

12b. DISTRIBUTION CODE

13. ABSTRACT (maximum 200 words)

Today wireless communications offer great convenience and efficiency, but concerns about security must be addressed. Wireless communications are considered less secure than wired or fiber-based systems because the data is transmitted over the radio channel making it more susceptible to eavesdropping and interception. Thus, security needs special attention. Confidentiality, integrity and availability are the objectives of security solutions. Attacks such as Man-in-the-Middle, replay, and Denial-of-Service are mitigated or eliminated by solutions such as those discussed in this thesis. Data disclosure to unauthorized people, user identity and location disclosure, impersonation of a valid user, user tracking and subscriber capabilities disclosure are a few of the potential risks that can lead to a mission failure and even cost people's lives.

This thesis explores how to securely leverage three cellular technologies, 3G, 4G/LTE and WiMAX, through an analysis of their security features. The security architectures of these wireless technologies are described. Their security vulnerabilities and the potential attack vectors are analyzed. A few protocols and techniques that address or mitigate the security deficiencies and the way they enforce security are provided. Furthermore, the importance of security in military communications is considered.

14. SUBJECT TERMS 3G,4G,LTE,WiMAX,Wireless Security ,Wireless Communications, Military Communications	**15. NUMBER OF PAGES** 115
	16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	**18. SECURITY CLASSIFICATION OF THIS PAGE**	**19. SECURITY CLASSIFICATION OF ABSTRACT**	**20. LIMITATION OF ABSTRACT**
Unclassified	Unclassified	Unclassified	UU

NSN 7540-01-280-5500

Standard Form 298 (Rev. 2-89)
Prescribed by ANSI Std. 239-18

THIS PAGE INTENTIONALLY LEFT BLANK

Approved for public release; distribution is unlimited

SECURE MILITARY COMMUNICATIONS ON 3G, 4G AND WIMAX

Panagiotis Schoinas
Lieutenant, Hellenic Navy
B.S., Hellenic Naval Academy, 2003

Submitted in partial fulfillment of the
requirements for the degree of

MASTER OF SCIENCE IN COMPUTER SCIENCE

from the

NAVAL POSTGRADUATE SCHOOL
September 2013

Author: Panagiotis Schoinas

Approved by: Gurminder Singh
Thesis Advisor

John H. Gibson
Co-Advisor

Peter J. Denning
Chair, Department of Computer Science

THIS PAGE INTENTIONALLY LEFT BLANK

ABSTRACT

Today wireless communications offer great convenience and efficiency, but concerns about security must be addressed. Wireless communications are considered less secure than wired or fiber-based systems because the data is transmitted over the radio channel making it more susceptible to eavesdropping and interception. Thus, security needs special attention. Confidentiality, integrity and availability are the objectives of security solutions. Attacks such as Man-in-the-Middle, replay, and Denial-of-Service are mitigated or eliminated by solutions such as those discussed in this thesis. Data disclosure to unauthorized people, user identity and location disclosure, impersonation of a valid user, user tracking and subscriber capabilities disclosure are a few of the potential risks that can lead to a mission failure and even cost people's lives.

This thesis explores how to securely leverage three cellular technologies, 3G, 4G/LTE and WiMAX, through an analysis of their security features. The security architectures of these wireless technologies are described. Their security vulnerabilities and the potential attack vectors are analyzed. A few protocols and techniques that address or mitigate the security deficiencies and the way they enforce security are provided. Furthermore, the importance of security in military communications is considered.

THIS PAGE INTENTIONALLY LEFT BLANK

TABLE OF CONTENTS

THIS PAGE INTENTIONALLY LEFT BLANK

LIST OF FIGURES

THIS PAGE INTENTIONALLY LEFT BLANK

LIST OF TABLES

THIS PAGE INTENTIONALLY LEFT BLANK

LIST OF ACRONYMS AND ABBREVIATIONS

3G	Third Generation
3GPP	Third Generation Partnership Project
3GPP2	Third Generation Partnership Project 2
AAA	Authentication Authorization and Accounting
ACC	Accumulator
AES	Advanced Encryption Standard
AGW	Access Gateway
AK	Anonymity Key (in 3G security)
AK	Authorization Key (in 4G security
AKA	Authentication and Key Agreement
AMF	Authentication Management Field
AS	Access Stratum
ASN	Access Service Network
AuC	Authentication Center
AUTN	AUthentication TokeN
AUTS	re-synchronization Authentication token
AV	Authentication Vector
BS	Base Station
BSS	Base Station Subsystem
CK	Cipher Key
COTS	Commercial Off The Shelf
C-RNTI	Cell Radio Network Temporary Identifier
CSIM	CDMA Subscriber Identity Module
CSN	Connectivity Service Network
DH	Diffie-Hellman protocol
DK	Delegation Key
EAP	Extensible Authentication Protocol
EAP-AKA	Extensible Authentication Protocol - Authentication and Key Agreement
EAP-TLS	Transport Layer Security

EAP-TTLS MS-CHAP v2	Tunneled TLS Tunneled Transport Layer Security with Microsoft Challenge-Handshake Authentication Protocol version 2
EC-AKA2	Ensured Confidentiality Authentication and Key Agreement II
EMSK	Enhanced Master Session Key
EMSUCU	Enhancement Mobile Security and User Confidentiality for UMTS
eNb	eNodeB
EPC	Evolved Packet Core
ePDG	evolved Packet Data Gateway
EPS	Evolved Packet System
EPS- AKA	Evolved Packet System- Authentication and Key Agreement
ESP	Encapsulating Security Payload
E-UTRAN	Evolved UMTS Terrestrial Radio Access Network
GGSN	Gateway GPRS Support Node
GKEK	Group Key Encryption Key
GMSC	Gateway Mobile Switching Center
GPRS	General Packet Radio Service
GSM	Global System for Mobile Communication
GTEK	Group Traffic Encryption Key
GUTI	Globally Unique Temporary Identity
HA	Home Agent
HE	Home Environment
HLR	Home Location Register
HLR/AuC	Home Location Register /Authentication Center
HMAC	Hashed Message Authentication Code
HSS	Home Subscriber Server
IK	Integrity Key
IKE	Internet Key Exchange
IMEI	International Mobile Equipment Identity
IMSI	International Mobile Subscriber Identity

IPsec	Internet Protocol security
KAC	Key Administration Centers
K_{ASME}	Key Access Security Management Entity
KEK	Key Encryption Key
K_{eNB}	Key eNodeB
K_{NASenc}	Key Non Access Stratum encryption
K_{NASint}	Key Non Access Stratum integrity
KSI_{ASME}	Key Set Identifier Access Security Management Entity
K_{UPenc}	Key User Plane encryption
LAI	Location Area Identity
LTE	Long Term Evolution
MAP	Mobile Application Part
ME	Mobile Equipment
MME	Mobile Management Entity
MS	Mobile Station
MSC	Mobile Switching Center
MSK	Master Session Key
NAP	Network Access Provider
NAS	Non Access Stratum
NDS	Network Domain Security
NE	Network Elements
NSP	Network Service Provider
PEK	Permanent pre-shared Encryption Key
PGW	Packet Data Network Gateway
PIK	Permanent pre-shared Integrity Key
PK_H	Public Key of HSS
PKMv2	Privacy and Key Management version 2
PLK	Payload encryption Key
PMK	Pairwise Master Key
P-TMSI	Temporary Mobile Subscriber Identity in Packet Switched (PS) Domain
RADIUS	Remote Authentication Dial In User Service

RANAP	Radio Access Network Application Part
RAND	Random challenge
RNC	Radio Network Controller
RRC	Radio Resource Control
RRM	Radio Resource Management
RSA	Rivest Shamir Adleman
SA	Security Association
SAE	System Architecture Evolution
SAID	Security Association Identity
SA-TEK	Security Association Traffic Encryption Key
SBC	Subscriber station Basic Capabilities
SE EPS- AKA	Security Enhanced Evolved Packet System-Authentication and Key Agreement
SEG	Security Gateway
SGPRS	Serving General Packet Radio Service
SGSN	Serving GPRS Support Node
SGW	Serving Gateway
SIM	Subscriber Identity Module
SK_H	HSS private key
SK_U	UE private key
SNID	Service Network Identity
SQN	Sequence number
SQN_{HE}	Sequence number counter maintained in the HLR/ AuC
SQN_{MS}	Sequence number counter maintained in the USIM
SRNC	Serving Radio Network Controller
SS	Subscriber Station
SS7	Signaling System No 7
S-TMSI	System Architecture Evolution - Temporary Mobile Subscriber Identity
TCP/IP	Transmission Control Protocol/Internet Protocol
TK	Temporary Key
TLS	Transport Layer Security

TMSI	Temporary Mobile Subscriber Identity in Circuit Switched(CS) Domain
UE	User Equipment
UESecCapabilities	User Equipment Security Capabilities
UMTS	Universal Mobile Telephone System
USIM	Universal Subscriber Identity Module
VLR	Visited Location Register
VLR/SGSN	Visited Location Register Serving General Packet Radio Service Support Node
WAP	Wireless Application Protocol
WAP 2.0	Wireless Application Protocol 2
WPKI	Wireless Public Key Infrastructure
WTLS	Wireless Transport Layer Security
XRES	expected response

THIS PAGE INTENTIONALLY LEFT BLANK

ACKNOWLEDGMENTS

I would like to thank my thesis advisors, Gurminder Singh and John H. Gibson, for their guidance and support in completing this thesis.

Finally, I would like to thank my family and my Greek and international friends for their support during my studies.

THIS PAGE INTENTIONALLY LEFT BLANK

I. INTRODUCTION

A. PROBLEM DESCRIPTION

The popularity and availability of wireless communications, particularly cellular, continues to grow rapidly world-wide. Mobile users are interested in services such as mobile shopping, mobile banking and mobile payments. Multimedia applications, high data rate, mobility, and cost make wireless communication one of the most useful means of communication. In the military, wireless communications support mobility and provide flexibility for troops on the battlefield or ships at sea, enabling them to communicate with command elements or higher headquarters.

Even though wireless communications offer great convenience and efficiency, concerns about security must be addressed. Wireless communications are considered less secure than wired or fiber-based systems because the data is transmitted over the radio channel making it more susceptible to eavesdropping and interception. Thus, security needs special attention. Confidentiality, integrity and availability are the objectives of security solutions. Attacks such as Man-in-the-Middle, replay, and Denial-of-Service are mitigated or eliminated by solutions such as those discussed in this thesis. Data disclosure to unauthorized people, user identity and location disclosure, impersonation of a valid user, user tracking and subscriber capabilities disclosure are a few of the potential risks that can lead to a mission failure and even cost people's lives.

This thesis explores how to securely leverage three cellular technologies through an analysis of their security features. The security architectures of these wireless technologies are described. Their security vulnerabilities and the potential attack vectors are analyzed. A few protocols and techniques that address or mitigate the security deficiencies and the way they enforce security are provided. Furthermore, the importance of security in military communications is considered.

B. THESIS OBJECTIVES

The main question that is addressed in this research is: “What are the improvements needed in order to securely leverage the 3G, 4G/LTE, and WiMAX cellular communications?”

Corollary questions to be answered in pursuit of this question are as follows:

- What security challenges do these wireless technologies pose?
- What potential security-related attacks could be mounted against these wireless technologies?
- What is the impact of the security vulnerabilities for military communications?
- In what ways may attackers be prevented from causing harm to communications by the solutions suggested?

For each technology, the report provides background information pertinent to the specific security architecture and associated vulnerabilities. Then solutions that address these vulnerabilities are discussed and analyzed. Finally, the potential impact of the vulnerabilities on military applications is analyzed and the benefits of applying the suggested solutions are presented.

C. THESIS OUTLINE

The thesis is organized into the following chapters:

- Chapter II describes the basic concepts and terminology and provides the necessary theoretical background by analyzing the security architectures of the three technologies.
- Chapter III describes the vulnerabilities and potential attacks that may be mounted, the suggested solutions and their method of enforcing security for these technologies.
- Chapter IV analyzes the suggested solutions and the security challenges they address, as well as provides the security benefits of security measures for military communications.
- Chapter V summarizes the analytical results and makes recommendations for future work.

II. BACKGROUND AND CURRENT CAPABILITIES

A. OVERVIEW OF RELEVANT WIRELESS NETWORKING TECHNOLOGIES

When talking about security in a mobile system, a few objectives come to mind. First, one must ensure that only legitimate users have access to the mobile system. Second, effort must be made to maintain a user's or operator's data confidentiality and integrity. Third, protection from denial of service (that is, assurance of user access) must be provided. Finally in the event of loss or theft of end-user devices, remote access by the administrators must be assured to maintain the mobile system's security [1].

This chapter provides information regarding the security architecture of Third Generation (3G), Long Term Evolution (LTE), and WiMAX cellular systems cellular systems. The background information will serve as a basis for the consideration of security threats and vulnerabilities of these wireless technologies.

1. UMTS Security

The Third Generation (3G) proposal for cellular communication aimed at maintaining compatibility with Global System for Mobile Communication (GSM) as well as address security weaknesses of the GSM architecture. Some of 3G's main security objectives include the following [2]:

- Ensure that information generated by the user is protected against misuse or misappropriation.
- Ensure that resources and services are protected adequately against misuse or misappropriation.
- Ensure that the security features are globally compatible.
- Protect the users in cases of stolen mobile stations or misused by monitoring their traffic and keeping track of mobile stations' identities.

- Use shared symmetric key for challenge and response messages between the SIM card and the authentication center during authentication procedure.
- Use unique user numbering, identification, and equipment during authentication.
- Ensure that the security features can be extended and applied to new services.

The Universal Mobile Telephone System (UMTS) security architecture, which is intended as a framework for implementing the previously stated objectives, is depicted in the following diagram. It is composed of five distinct security features/classes, as enumerated, which address specific threats and provide specific protection mechanisms for each threat. The key aspects of the five classes are discussed in the remainder of this section [1], [3]:

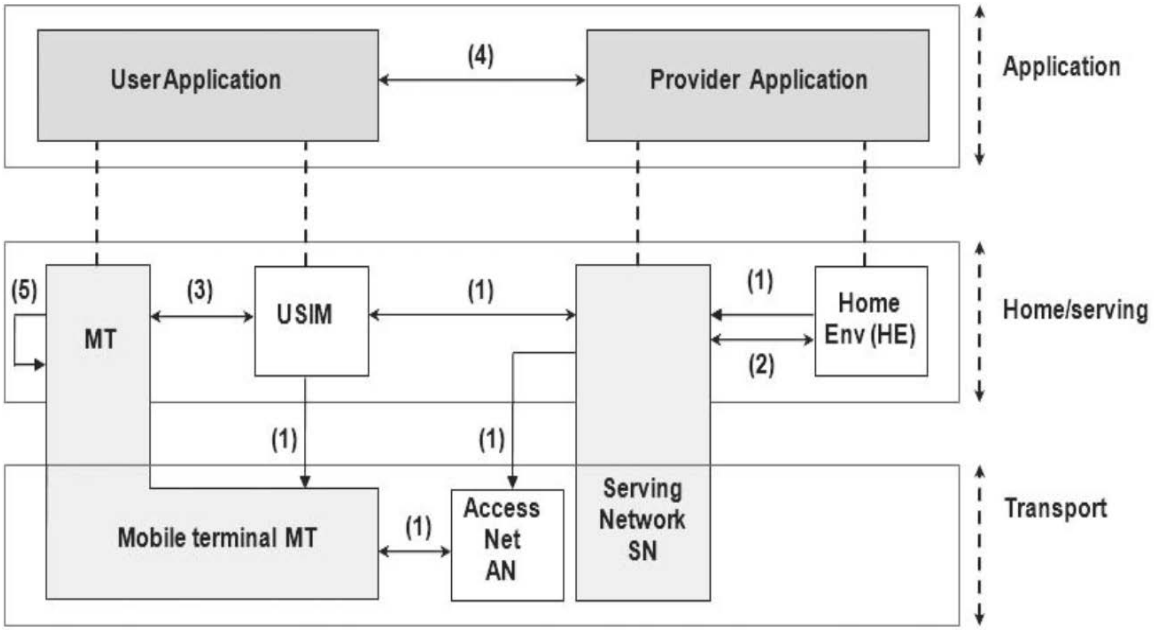

Figure 1. UMTS security architecture, from [3]

- Network access security (class 1): enables the user to securely access a 3G network and provide protection from attacks on the (radio) access link.

- Network domain security (class 2): enables nodes in the provider domain to securely exchange signaling messages and provide protection from attacks on the wire-line network.
- User domain security (class 3): allows only authorized access to mobile terminals.
- Application domain security (class 4): ensures the secure message exchange between user and provider application domains.
- Visibility and configurability of security (class 5): enables the user to be informed about which security features are in operation (and which are not) and which services are based on the security features.

a. Network Access Security

Network Access Security includes entity authentication, confidentiality, and data integrity functions [1, 3]. These functions refer to user identity confidentiality, authentication and key agreement and data confidentiality and integrity protection of signaling messages.

(1) User Identity and Location Confidentiality. The user identity confidentiality feature prevents a user's information and location disclosure. It specifically impedes passive user data eavesdropping to protect the user's identity. The user is assigned a Temporary Mobile Subscriber Identity (TMSI/P-TMSI) and is identified by the TMSI on the radio access link, except during the user's first registration, where the TMSI is not generated until the user is verified by the permanent identity. The visited Location Register is responsible for tracking the mapping between the permanent (International Mobile Subscriber Identity (IMSI)) and temporary (TMSI) identity. Whenever a user changes a location the temporary identity is acquired from the previous VLR if possible; otherwise, the same procedure of permanent identity request is followed. Moreover, in order to avoid the compromise of a user's identity and location, the temporary identity assigned to the user changes after a period of time, making it difficult for penetrators to track the user. Furthermore, any signaling message or user packet that may include information on user identity is encrypted by the radio interface.

Thus, in order to achieve user identity confidentiality, user location confidentiality and user intractability, a few features are used that are identical to those used on the GSM networks. First, the serving network enforces identification of the mobile equipment by requesting from terminal the mobile equipment's IMEI (International Mobile Equipment Identity). A paradox is that the mobile IMEI cannot be verified because it is based on the terminal's legitimacy, which provides the IMEI. Secondly, only authorized users have access to the Universal Subscriber Identity Module (USIM). The USIM is a memory component that stores subscriber information and customs settings on a SIM card. The authentication is enforced by using a shared secret key (personal identity number) that is stored securely in USIM; the user has to know it in order to authenticate. Lastly, a shared secret key between USIM and mobile terminal ensures that only the authorized USIM has the ability to access the mobile terminal [1, 3].

(2) Authentication and Key Agreement. The UMTS authentication and key agreement (AKA) [1, 3] achieves mutual authentication between the user and the serving network as well as the establishment of a cipher and integrity key. The authentication takes place in the USIM on the user side and in the HLR/AuC on the network side. The mechanism of AKA is depicted in the Figure 2.

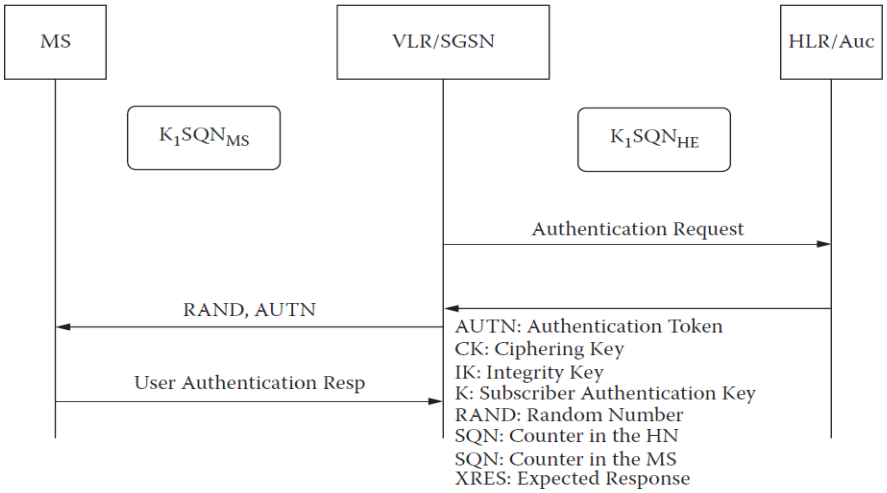

Figure 2. Authentication and key agreement, from [3].

Based on a challenge/response authentication protocol, the two involved parties do not reveal or transmit their secret password but use it in order to confirm the other party's identity. The USIM uses a sequence counter, SQN_{MS}, to show the highest sequence number the USIM has accepted during the network authentication procedure, and the home equipment uses another value, SQN_{HE}, a unique sequence number for each individual user. Once VLR/SGSN makes an authentication request, the HLR/AuC responds by sending an array of *"n"* authentication vectors, which are ordered according to the sequence number. A random number, RAND, an expected response, XRES, a cipher key, CK, an integrity key, IK, and an authentication token, AUTN, comprise every authentication vector.

The way the authentication vectors are generated is depicted in Figure 3.

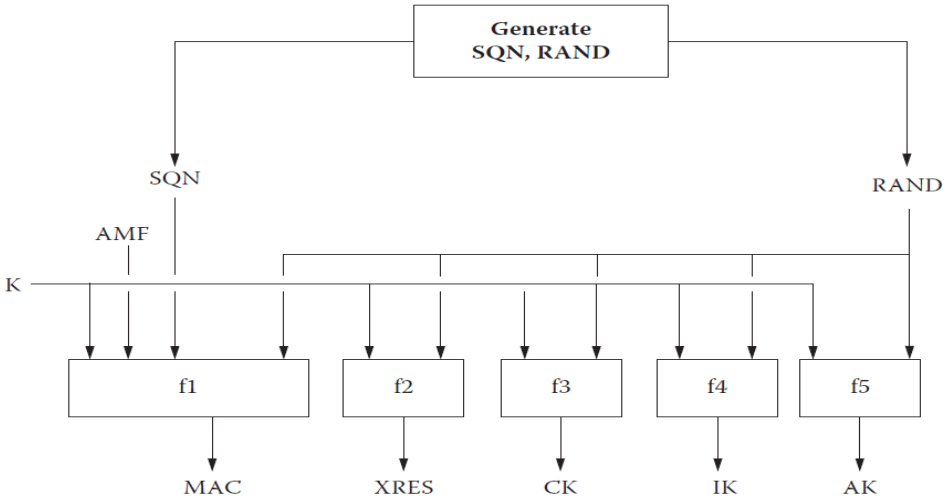

Figure 3. Generation of authentication vectors, from [3].

The HLR/AuC generates a fresh (not previously used) sequence number, SQN, and a random unpredictable challenge value, RAND. Then, the rest of the components that compose the authentication vector are produced by using a set of functions, *f1* through *f5* that are described in detail in [4],, and the secret key, *K*, as follows [3]:

- using a message authentication function, *f1,* it computes the message authentication code, MAC:

  MAC = f1 (K, SQN, RAND, AMF)

  where AMF is the authentication and key management field that is used to improve the performance or bring a new authentication key into use.
- using a (possibly truncated) message authentication function, *f2,* it computes the expected response that is going to be compared later in the VLR/SGSN with the response received from MS:

  XRES= f2 (K, RAND)
- using a key generating function, *f3,* it computes the cipher key, CK:

  CK = f3 (K, RAND)
- using a key generating function, *f4,* it computes the integrity key, IK:

  IK = f4 (K, RAND)

- using a key generating function, *f5*, it computes the anonymity key, AK:

  AK = f5 (K, RAND)

- assembles the authentication token, AUTN, and updates the counter, SQN_{HE}:

  AUTN = <SQN ⊕ AK, AMF, MAC>

After all these components have been specified, the VLR/SGSN forwards parameters, RAND and AUTN, to the MS. The USIM, using the same secret key, K, computes the anonymity key,

AK = f5 (K, RAND),

and retrieves the SQN,

SQN = (SQN ⊕ AK) xor AK.

Then, USIM computes the expected message authentication code, XMAC,

XMAC = f1 (K, SQN, RAND, AMF),

and compares it with the MAC that is included in AUTN. If the AUTN and the SQN are accepted then USIM computes the response, RES,

RES= f2 (K, RAND)

and sends it to the VLR indicating a successful receipt. If MAC and XMAC are not the same, the MS abandons the procedure by sending back an authentication response message and stating that there is an integrity failure. If SQN is not acceptable, the MS abandons the procedure by sending back a synchronization failure response message and the computed re-synchronization token, AUTS.

At the same time, the USIM computes the CK and IK,

CK = f3 (K, RAND)

IK = f4 (K, RAND).

Finally, the VLR/SGSN compares the RES with the XRES and if they match the AKA is successful and it forwards the CK and IK to the ME and the corresponding radio network controller (RNC) to enforce encryption and integrity.

(3) Integrity Protection of Signaling Messages. This integrity protection mechanism of signaling messages [1, 3, 5] is depicted in the following figure:

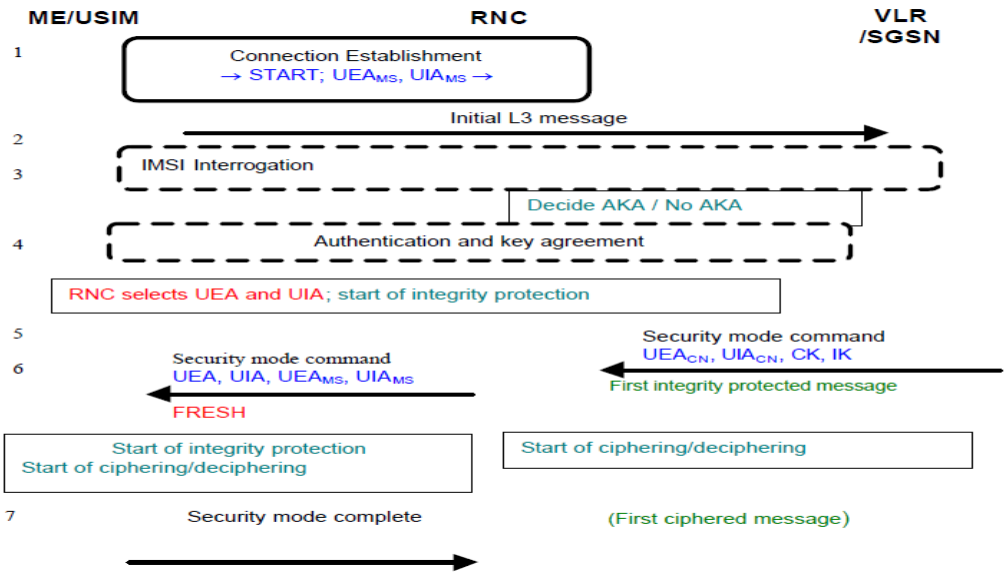

Figure 4. Starting ciphering and integrity, from [1].

During Radio Resource Control (RRC) establishment, the ME sends the START values for the CS and PS domain and the UMTS Encryption algorithms (UEAMS) and integrity algorithms (UIAMS) to the RNC, informing the RNC regarding the ciphering and integrity capabilities of the ME. Then, the MS sends the Initial Layer 3 L3 message (location update request, routing area update request, paging response, etc.) to the VLR/SGSN, including the user identity and the Key Set Identifier (KSI), which is the one used during the last authentication. There may be an IMSI interrogation, user authentication, CK and IK production and a new KSI allocation required. After VLR/SGSN specifies the allowed UEA and UIA to use, it starts the integrity and ciphering protection by sending the Radio Access Network Application Part (RANAP) message, called the Security Mode command, to the Serving Radio Network

Controller (SRNC).This is the first integrity protected message and contains the UIA_{CN}, the IK (the UEA_{CN} and the CK, if ciphering is starting, too). If a new key generation takes place, the START value should be reset. The SRNC selects the algorithms to use by checking from both the allowed and the MS capable algorithms, generates a FRESH value and starts the integrity protection. If there is a conflict with the data received from VLR/SGSN, then a security mode reject message is sent back. Thereafter, the SRNC computes the expected MAC and includes it with the UEAs and UIAs in the RRC message Security Mode command that is sent to the ME. Then the MS checks all the received security capabilities that are the common with those included in the initial message and generates and compares the MAC with the XMAC that it received. If all checks are satisfied a response message that Security Mode is complete is sent to SRNC, which in turn verifies the integrity and forwards it to the VLR/SGSN, including the selected algorithms [5].

In this manner, the integrity protection mechanism of signaling messages is enforced and does not permit malevolent entities to hijack the connection or spoof a message.

(4) Signaling Data Integrity Mechanism. The function algorithm, f9, which is described in detail in [6] is used to protect against false base station attacks so that the receiving entity (MS or SN) is able to verify the message's originality and non-modification (integrity). This algorithm implements the KASUMI algorithm and is based on a chain of block ciphers, whose 64-bit output is used to generate the 32-bit Message Authentication Code [1].

The verification process takes place in the ME and in the RNC. At first, in the sender side a 32-bit MAC is computed based on the f9 algorithm. In addition to the 128-bit integrity key, IK, and the variable length frame (MESSAGE), the f9 algorithm uses a time dependent 32-bit value, COUNT, a randomly generated 32-bit value, FRESH, and a 1-bit value, DIRECTION (showing the direction of transmission), to differentiate two identical messages [1]:

MAC = f9 (IK, MESSAGE, COUNT, FRESH, DIRECTION)

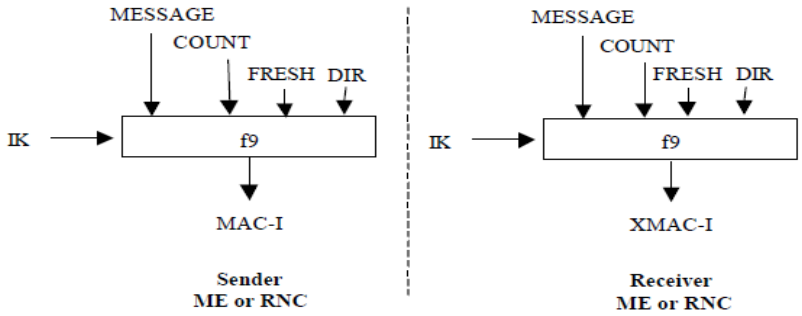

Figure 5. Air interface integrity mechanism, from [1].

The integrity check is completed in four steps [1]:

- The f9 function computes the MAC based on the inputs and the MESSAGE.
- The MAC is attached with the signal and sent to receiver.
- The receiver computes the expected XMAC.
- From the receiver side the integrity check is completed by the comparison of the received MAC and the calculated XMAC.

(5) Air Interface Confidentiality Mechanism. The function algorithm, f8, which is described in detail in [6], is used to protect the user and signaling data that are exchanged between RNC and MS over the radio interface. The algorithm is a symmetric synchronous stream cipher that is used for encryption on the sender side as well as decryption on the receiver side. It is based on the KASUMI algorithm, which is analyzed in [7]. In addition to the 128-bit cipher key, CK, and the variable length frame, called *length*, the f8 algorithm uses a time dependent 32-bit value, COUNT, a 5-bit bearer identity value, BEARER, and a 1-bit value, DIRECTION (showing the direction of transmission) to generate the output keystream block. The keystream has the same length as the original frame. The cipher-text is the result of a bitwise XOR operation between the plaintext and the keystream.

keystream = f8(CK, BEARER, DIRECTION, length)

Cipher-text = keystream ⊕ plaintext

The confidentiality mechanism consists of the following four steps [1]:

- The f8 function computes the output keystream based on the inputs identified earlier.
- The keystream is XORed with the plaintext to generate the resulting cipher-text.
- The cipher-text is sent to the receiver through the radio interface.
- On the receiver side, the keystream is computed the same way as on the sender side and by applying the bitwise XOR operation between it and the received cipher-text the plaintext is extracted.

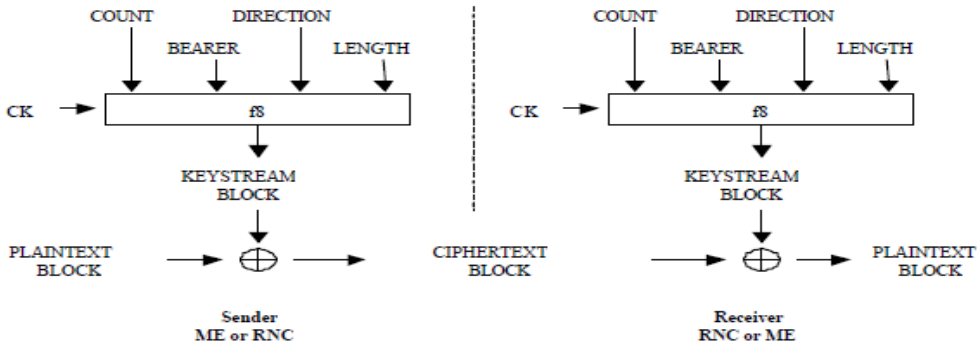

Figure 6. Air Interface confidentiality mechanism, from [1].

b. Network Domain Security

The Network Domain Security (NDS) [1, 3] offers security between entities that may be in the same or different networks. There are many different protocols and interfaces that may be used for network security, like Mobile Application Part (MAP) [8] and General Packet Radio Service tunneling (GPRS). Internet Protocol based security protocols (IPsec-based), which are applied at the network layer, and Signaling System No. 7 (SS7-based) protocols, which are

applied at the application layer, are based on the existing cryptographic techniques [1].

(1) IP-based protocols. In a UMTS network, usually the operators are related with a security domain. The security gateways (SEGs) are the entities at the border of the IP security domain that protect the IP-based protocols and all the traffic utilizing them. All the security domains make up the network domain security control plane, which is restricted to the network domain and does not provide protection to the user plane [1, 3].

The IP Security Association (SA) negotiations that take place between the Key Administration Centers (KAC) on behalf of the entities and security gateways are enforced through the Internet Key Exchange (IKE) protocol and their distribution through standard interfaces. The 3GPP determined that Encapsulating Security Payload (ESP) should always be used for protection of the packets, which enforces both confidentiality and integrity protection, and Advanced Encryption Standard (AES) should be used as an encryption algorithm in internetworking solutions. For node authentication there are two options: pre-shared symmetric keys or public key infrastructure. Lastly, the IPsec can be configured in transport and tunnel mode. The tunnel is preferred when one of the entities is a security gateway [1, 3].

There are two modes of enforcing the IP security [1]:

- Hop by hop: distinct IPsec tunnels are established between every pair of network entities. That means the route that a message has to follow between two entities that are in different networks is the following: the sending entity establishes an IPsec tunnel to the SEG in its security domain and then forwards the data. Then the SEG terminates the tunnel, establishes a new one with the appropriate SEG in the network where the receiving entity belongs and sends the data. The receiving SEG now terminates the previous tunnel, establishing a new one with the receiving entity and forwards the data
- end to end: in this mode the security association takes place between the sending and receiving entity, and it is applied even for entities within the same network.

(2) SS7- based protocols. When the traffic transport is based on the SS7 protocol only, or on a combination of SS7 and IPsec protocols, the security is enforced in the application layer. However, when it is based on IPsec only, the security is enforced either at the network layer only or in addition to the application layer. The mobile specific part of SS7 signaling is the MAP, and the complete set of procedures that enforce security for MAP messages is called the MAPsec. The MAPsec provides security by encrypting the original message and putting it in another MAP message and using a message authentication code generated for the original message and added in the new MAP message. The security association procedure at the application layer is network-based, similar to the one that was previously described for IP-based protocols. The SAs contain cryptographic keys and KACs are in charge of SA negotiations and distributions. Furthermore, the end-to-end solution enforces non-disclosure to entities other than the sender and the receiver [1, 3].

c. User Domain Security

User Domain Security [1, 2, 3] consists of mechanisms that enforce secure access to mobile stations. The security relies on a removable card, called the UMTS integrated security card, and security applications, like USIM, CSIM (CDMA Subscriber Identity Module), or SIM, which all execute on this card. The USIM is the module through which user identification and association to home equipment is enforced, as noted earlier. For 3G networks, USIM is in charge of key agreement, as well as subscriber and network authentication.

The User domain security contains two types of authentications: the user-to-USIM authentication and USIM-to-terminal authentication. The USIM gives access to a user (or users) after the user proves the knowledge of a shared secret key stored in the USIM. Moreover, only an authorized user is allowed to have access to other user equipment or to a terminal. A secret key that is shared between the USIM and the terminal and stored securely in both entities is used to enforce access control [2], [3].

d. Application Domain Security

The Application Domain Security [1, 2, 3] refers to the security of a message's exchange between the mobile station and the serving network or service provider, whereas the network operator or the application provider chooses the security level. A user is allowed to use applications only after being authenticated. At the same time, application level confidentiality could also be enforced. Since USIM gives the opportunity to operators or third party providers to create an application, the secure exchange of messages should be ensured. This is achieved through numerous security mechanisms [2, 3]:

- entity authentication of applications
- message authentication
- replay detection of application data
- sequence integrity of application data
- data integrity of application data
- confidentiality assurance
- proof of receipt

These mechanisms are assigned and incorporated in the USIM Application Toolkit. The USIM Application Toolkit is in charge of the applications' creation, which are resident on the USIM. All these security mechanisms at the application level are necessary so that protection is enforced, even if there is no end-to-end security mechanism enforced in lower layers.

Two of the most popular application protocols, Wireless Application Protocol (WAP) and Wireless Application Protocol 2 (WAP 2.0) that include a set of standards for accessing information over mobile wireless networks, use two different mechanisms to achieve security in the communication.

In WAP architecture, the WAP gateway translates the protocols used in the WAP segment to the protocols used in the public Internet, enabling the connection between the wireless domain and public Internet. As regards security, the Wireless Transport Layer Security (WTLS) protocol is used. Since it supports datagrams in low-bandwidth/high-latency environments, WTLS provides

an optimized handshake through dynamic key refreshing, which allows encryption keys to be regularly updated during a secure session. Thus, it is used to enforce data integrity, privacy, authentication and denial of service (DoS) protection. Therefore, the WAP gateway is in charge of managing wireless security and carrying secure data between WTLS and TLS security channels for Web applications that follow public Internet standards with TLS [2, 3.]

In WAP2 architecture, the introduction of existing IP-stack protocols into the WAP-environment was the major difference from the original WAP architecture. This way, many different gateways are allowed, and the conversion between the protocol stacks is available anywhere. A TCP-level gateway allows wired and wireless versions of TCP respectively, and a Transport Layer Security TLS channel established between the mobile device, and the server runs on top of TCP. The wireless profile of TLS includes many security features, like cipher suites, signing algorithms, certificate formats and the use of session “resume.” Thus, taking advantage of these benefits, an end-to-end security capability is enforced at the transport level, making secure communication feasible [2, 3].

e. Visibility and Configuration of Security

The features that belong to this class inform the users about which security features are effective (and which are not), as well as on which security features (if any) the use of particular services depends. Thus, the user should be offered visibility into operation of the security features, like indication of the access network encryption, the network wide encryption and the level of security provided [2, 3].

Moreover, the configurability offers the user and HE the opportunity to configure required features depending upon the service provisioning needed. It is obligatory for all security features upon which a required service depends to be in operation in order for the service to be available. Some of the configurability features include: enabling/disabling user-USIM authentication,

accepting/rejecting incoming non-ciphered calls, establishing non-encrypted calls and accepting/rejecting certain encryption algorithms [2, 3].

2. LTE Security

LTE and WiMAX are both 4G wireless technologies. Relying on 3G, the designers' objectives for 4G were concentrated on improving performance. A few of their objectives were: high data rate, large number of simultaneously supportable users,, low cost per bit, low latency, good quality of service, good coverage and support for mobility at high speeds. Thus, 4G wireless technologies are based on 3G but with a few key differences, the most important of which is that they operate entirely based on the TCP/IP architecture and suite of protocols. For this reason, security issues arise since the technology is moved to an open set of communication protocols. To deepen our understanding of 4G networks, information about the network and security architecture of LTE and WiMAX technologies is provided in the rest of the chapter.

a. LTE Network Architecture

The LTE Network architecture [9, 10] is depicted in Figure 7. It relies on two basic Network Elements (NEs): the eNodeB (eNB), which is actually an improved base station, and the Access Gateway (AGW), which includes all the required functions for interfacing with the Evolved Packet Core (EPC). The eNB resides in the E-UTRAN (Evolved UMTS Terrestrial Radio Access Network). It is the entity by which the User Equipment (UE) is connected to the wireless network. The E-UTRAN connects to the EPC, which is IP-based, and the EPC subsequently connects to the provider wired IP network.

LTE incorporates several improvements over 3G. First, LTE only uses two basic NE types, the eNB and the AGW, as it does not require a circuit-switched interface while UMTS uses four basic elements the Mobile Switching Center (MSC), the Gateway Mobile Switching Center (GMSC), the Serving GPRS Support Node (SGSN) and the Gateway GPRS Support Node (GGSN). Second, the architecture is all IP-based, from the packet generated in the UE to

the signaling and control protocols used. Third, the architecture used is meshed, which improves the network's performance and offers reliability, efficiency and redundancy.

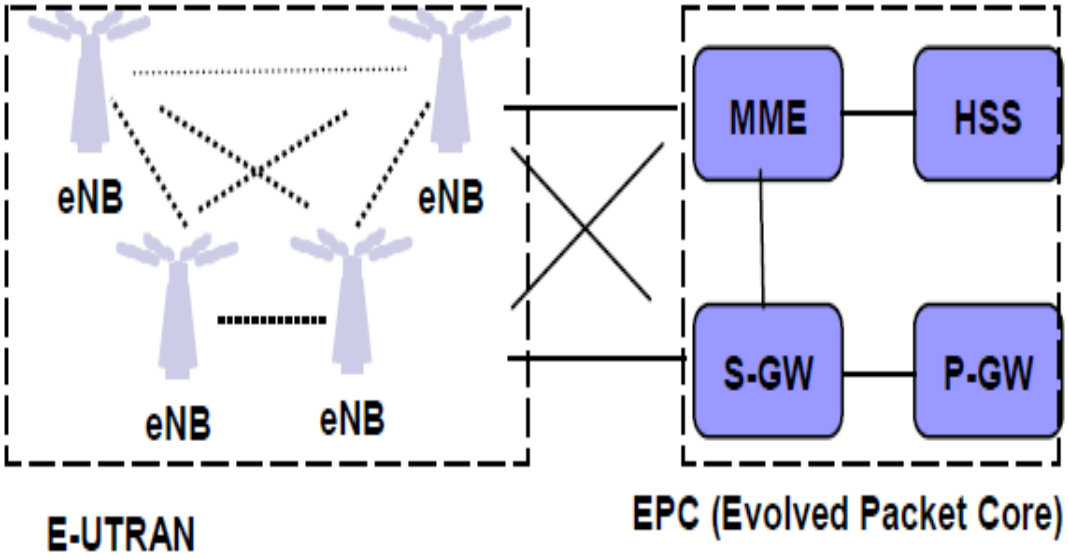

Figure 7. LTE system architecture evolution, from [9].

In order to understand how security is enhanced in the LTE network, a few more details about LTE components and how they cooperate is needed. Starting with the eNB, which is the single type of system in the E-UTRAN, it takes care of the radio-interface-related functions. Moreover, it is responsible for [9, 10]:

- inter-cell radio resource management (RRM) that coordinates resource allocation between different cell sites
- radio admission control that validates or disapproves the connection's establishment after a check is performed
- scheduling through dynamic resource allocation
- negotiation of QoS on the uplink
- compression/decompression of the packets that are transferred to or from the user equipment.

The second basic network entity in LTE, the AGW, includes the following modules [9, 10]:

- MME (Mobile Management Entity): the essential node for LTE. It is the node where mobility and security authentication, as well as management of user equipment identity, are accomplished. After the user equipment connects to the network, MME chooses the Serving Gateway. Moreover, after receiving the authentication data that was generated by HSS, it authenticates the user. Lastly, it is responsible for security key management and regulating user equipment roaming.
- HSS (Home Subscriber Server): handles the user information and security enforcement. All information needed for network entities to complete sessions that have to do with the user and the user's subscription reside in the HSS. The HSS generates the authentication data, which is later used by MME and is essential for the authentication and key agreement procedure between MME and user equipment. Instead of using SS7 to connect to the packet core (as was used in 3G), it relies on IP-based Diameter protocol, the authentication, authorization, and accounting protocol developed to succeed RADIUS. More information about RADIUS is provided in [11].
- SGW (Serving Gateway): is responsible for trafficking data packets. It terminates the interface towards E-UTRAN providing mobility to inter-eNodeB handovers as well as between LTE and other 3GPP technologies. Furthermore, it replicates the data packets, which is important in the case of legitimate interceptions.
- PGW (Packet Data Network Gateway): provides to the user equipment the option of connecting to more than one provider wired network since the user equipment (UE) can be connected to more than one PGW at the same time. This is the gateway through which UE connects to devices that do not reside in the service provider main IP network. A beneficial functionality of PGW is that it is essential for mobility between 3GPP and non 3GPP technologies. It is responsible for allocating the user equipment's IP address, as well as offering per user packet filtering, policy enforcement, and charging support.

Finally, the LTE network architecture enables these modules to work on the same or different devices due to its flexibility.

b. LTE Security Architecture

Since the follow-on to the initial cellular service (2G digital successor to the 1G analog service), wireless security has improved significantly. Based on already improved 3GPP security, LTE continues to make communications more secure. Measures have been taken to protect the user identity, secure signaling between the user equipment and MME, as well as secure communication between 3GPP networks and trusted non-3GPP users. Thus, compared to 3G, the key hierarchy and interworking security have improved. There are added security features for the eNodeB, and the authentication and key agreement was extended.

A schematic diagram of the LTE security overview is depicted in Figure 8:

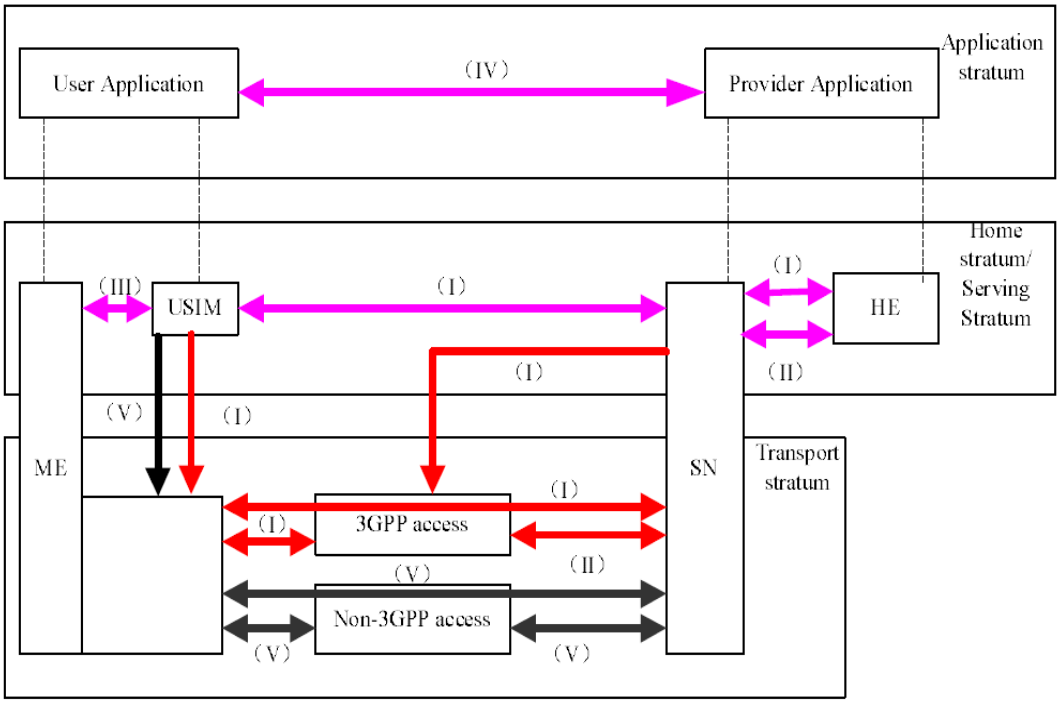

Figure 8. LTE security overview, from [10].

There are significant similarities between Figures 1 and 8 since LTE relies on the 3G architecture. The key difference is that instead of one Access Network, in LTE there are two distinct interfaces, one for 3GPP and one for non-3GPP access. For non-3GPP users the non-3GPP domain security (class V) depicted in the diagram allows the user equipment to access the EPC network securely through the non-3GPP network, enforcing at the same time radio access link protection [10].

The structure of the LTE security is concentrated on key security and hierarchy, authentication, encryption and integrity protection, key management, and user identity protection [10].

(1) Key Security and Hierarchy. A new key hierarchy is depicted in Figure 9 that enforces the security protection of signaling and user data traffic. In LTE, there are five distinct security-critical keys that are used for different purposes and have different life spans. Using the permanent key, K, that is stored on USIM, the ciphering, CK, and integrity, IK, keys are generated during an EPS (Evolved Packet System) AKA procedure, similar to that done for the 3G.Thereafter, the K_{ASME} (Access Security Management Entity) is generated based on CK, IK, and the SN identity. Subsequently, the integrity, K_{NASint}, and ciphering, K_{NASenc}, keys, used to provide security in NAS (Non Access Stratum) signaling messages between UE and MME, as well as the intermediate key, $K_{eNB,}$ are derived from $K_{ASME.}$ The K_{eNB} is derived in MME and UE; it is the one on which the keys, K_{UPenc}, K_{RRCint} and K_{RRCenc}, are based. The encryption key, K_{UPenc} (Key User Plane encryption), is used to protect the user messages exchanged between the UE and eNB. The integrity, K_{RRCint}, and encryption, K_{RRCenc}, keys provide RRC integrity and encryption protection, respectively, between the UE and eNB. A more detailed report about the five keys (K_{NASint}, K_{NASenc}, K_{UPenc}, K_{RRCint} and K_{RRCenc}), their length and purpose, as well as the intermediate keys from which they derive, are displayed in Table1.

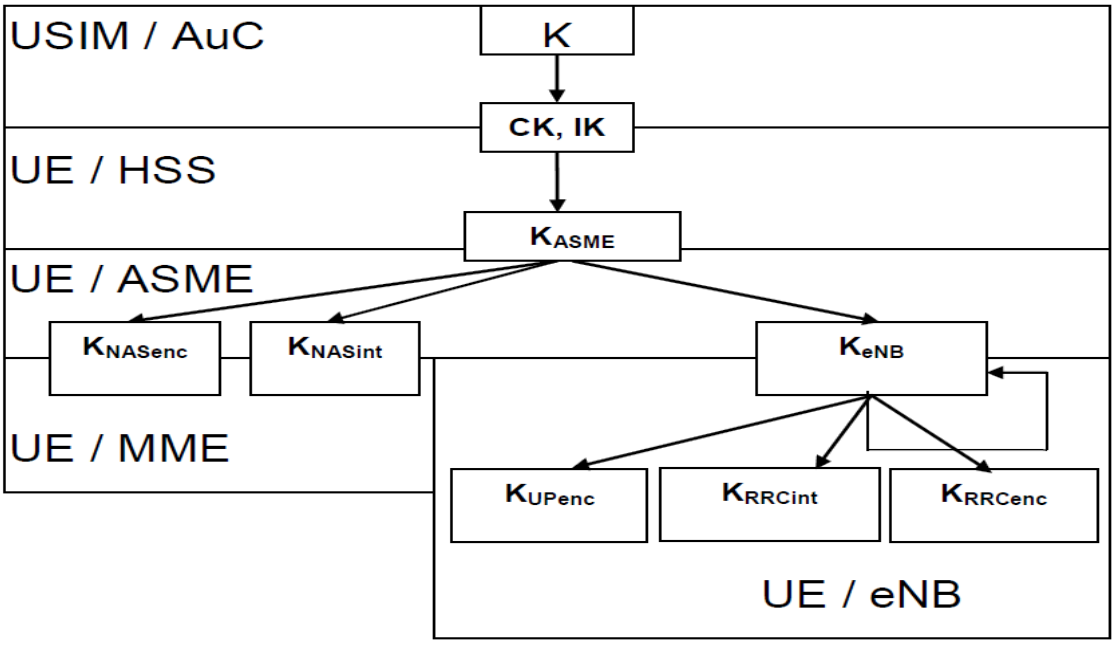

Figure 9. Key hierarchy of LTE, from [10].

Key	Purpose	Length	Derived from	Description
K	Master Base Key for GSM, UMTS, EPS	128	-	Secret Key stored permanently in USIM and AuC
CK, IK	Cipher and Integrity Keys	128	K	Pair of Keys derived in AuC and USIM during a AKA run. CK and IK should be handled differently for EPS as compared to Legacy context
K_{ASME}	MME (ASME) Base/Intermediate Key	256	CK, IK	Intermediate Key derived in HSS and UE from CK, IK during AKA. This is sent as part of the EPS AVs from HSS which include RAND, XRES, AUTN, and uniquely identified with eKSI allocated by the MME during AKA process. MME assumes the role of ASME in EPS
K_{eNB}	eNB Base Key	256	K_{ASME} K_{eNB}*	Intermediate Key derived in MME and UE from K_{ASME} when UE transits to ECM-CONNECTED State or by UE and Target eNB from K_{eNB}* during Handover
K_{eNB}*	eNB Handover Transition Key	256	K_{eNB} (H) NH (V)	Intermediate Key derived in Source eNB and UE during Handover when performing Horizontal (K_{eNB}) or Vertical Key (NH) Derivation. Used at Target eNB to derive K_{eNB}
NH	Next Hop	256	K_{eNB}	Intermediate Key derived in MME and UE used to provide forward security, and forwarded to eNB via the S1-MME interface
K_{NASint}	Integrity Key for NAS Signaling	256 (128 LSB)	K_{ASME}	Integrity Key for protection of NAS data derived in MME and UE
K_{NASenc}	Encryption Key for NAS Signaling	256 (128 LSB)	K_{ASME}	Encryption Key for protection of NAS data derived in MME and UE
K_{UPenc}	Encryption Key for User Plane (DRB)	256 (128 LSB)	K_{eNB}	Encryption Key for protection of user plane data derived in eNB and UE
K_{RRCint}	Integrity Key for RRC Signaling (SRB)	256(128 LSB)	K_{eNB}	Integrity Key for protection of RRC data derived in eNB and UE
K_{RRCenc}	Encryption Key for RRC Signaling (SRB)	256 (128 LSB)	K_{eNB}	Encryption Key for protection of RRC data derived in eNB and UE

Table 1. Summary description of EPS security keys, from [12].

(2) Authentication, Encryption and Integrity Protection. Fresh authentication vectors (AVs), good security algorithms, and use of IPSec are the anchors upon which the authentication, encryption, and integrity protection are based. First of all, the AVs are the vital elements of the authentication procedure, and the freshness is effected through the sequence numbers that are included in the exchanged messages. Freshness is a term used to state that the AVs are new and have not been used again. Freshness provides replay attack protection. Second, the security algorithms that are used in the HE and USIM to generate the authentication vectors are one-way mathematical functions making it difficult for the intruder to derive the input from the output. Third, the use of IPSec provides confidentiality to messages that are exchanged between nodes in the LTE EPS, as well as messages between nodes in home and visited networks [10].

(3) Key Management. LTE uses EPS AKA, depicted in Figure 10, for the authentication through which the keys are established and verified. EPS AKA starts with the UE sending its identity and continues with credentials exchange and challenge-response messages. In this way, security is provided for key management, which consists of key parameter establishment, generation and distribution.

(4) User Identity Protection. In order to prevent user identity disclosure to unauthorized entities, LTE minimizes the instances when the user's permanent identity is sent over the air by using temporary identifiers when possible. The various identifiers used, both temporary and permanent, are as follows [9], [10]:

- IMSI (International Mobile Subscriber Identity): A permanent identity that is sent in the clear when the associated user equipment attempts to initiate access to the network.
- IMEI (International Mobile Equipment Identity): A permanent identifier that is unique for every mobile device; used by companies to deny access when there is a report of a stolen mobile device, even if the SIM is replaced.
- M-TMSI (MME-associated Temporary Mobile Subscriber Identity): A temporary identifier that enforces subscriber confidentiality between UE and MME; the visited network assigns it after encryption. There is no feasible disclosure of the relationship between IMSI and M-TMSI to entities other than UE and MME.
- S-TMSI (System Architecture Evolved Temporary Mobile Subscriber Identity): A temporary identifier that is used for paging the UE; used by the network to request the establishment of a NAS signaling connection to the UE.
- GUTI (Globally Unique Temporary Identity): A temporary identifier that identifies uniquely the MME and the UE within the MME enforcing subscriber's confidentiality. It can also be used by the network and the UE during exchanging messages in order to establish user equipment identity in the EPS.
- C-RNTI (Cell Radio Network Temporary Identifier): A temporary identifier that is used to uniquely identify the UE at

the cell level and is assigned by the network. This is changed when UE moves to another cell.

There are more ways that LTE networks enforce security by using End-to-End Security. These include [9, 10]:

- Authentication and Key Agreement: a mutual authentication between the UE and the EPC takes place that is essential for LTE security through an AKA procedure. As depicted in Figure 10, authentication starts with the UE when it tries to connect to the EPC. The MME, representing the EPC, makes an authentication request to the HSS. The HSS that has the subscriber information verifies the authentication and generates authentication data that are later forwarded to the MME and verified by the UE. Moreover, during the AKA procedure more security keys are generated that are used for encryption and integrity protection. The procedure is similar to that for 3GPP access networks. The procedure is altered for non-3GPP access networks. First, the authentication takes place between the UE and the AAA (Authentication, Authorization, and Accounting) server, which resides in the EPC. The access authentication is based on the Extensible Authentication Protocol – AKA (EAP-AKA) procedure. The trusted non-3GPP access networks can be pre-configured in the UE, otherwise the non-3GPP access network is considered untrusted. In the case of an untrusted non-3GPP access network, an IPSec tunnel is established between the UE and the gateway ePDG (evolved Packet Data Gateway). This tunnel is used by the UE to pass the data through to the trusted ePDG that is connected to the EPC; it should rely on the Internet Key Exchange Protocol Version 2 (IKEv2) and the EAP-AKA.

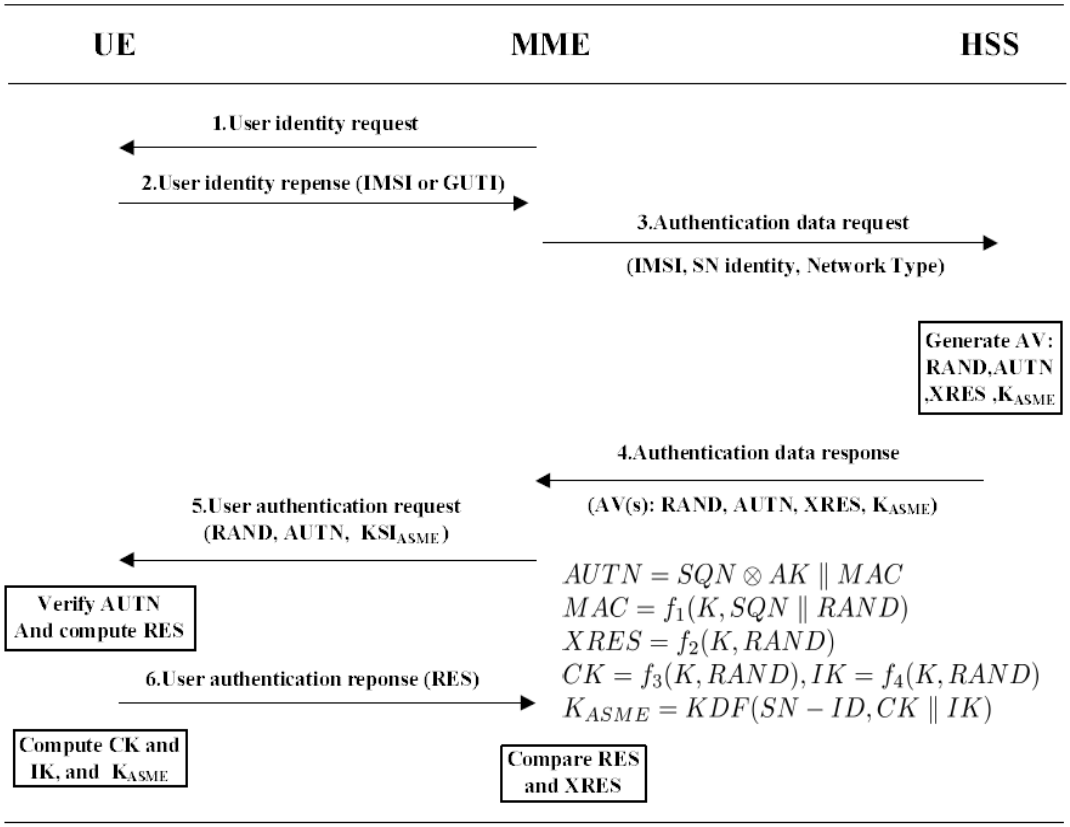

Figure 10. EPS AKA, from [10].

- Confidentiality and Integrity of Signaling: the RRC signaling between the UE and the eNB, as well as NAS signaling between the UE and the MME, are included in the Network Access Control plane. They are both encrypted providing confidentiality.
- User Plane Confidentiality: user plane data/voice is encrypted between the UE and the eNB to provide confidentiality. Moreover, in order to transport the user plane data, an IPSec tunnel can be established between the eNB and the SGW.

3. WiMAX Security

WiMAX consists of the wireless technologies that are based on IEEE 802.16 standards. In order to become familiar with WiMAX, information about WiMAX networks is provided so that the reader has an idea about some terms that are necessary to understand the security architecture.

a. WiMAX Network Architecture

In the figure below, a typical end-to-end WiMAX network architecture is depicted.

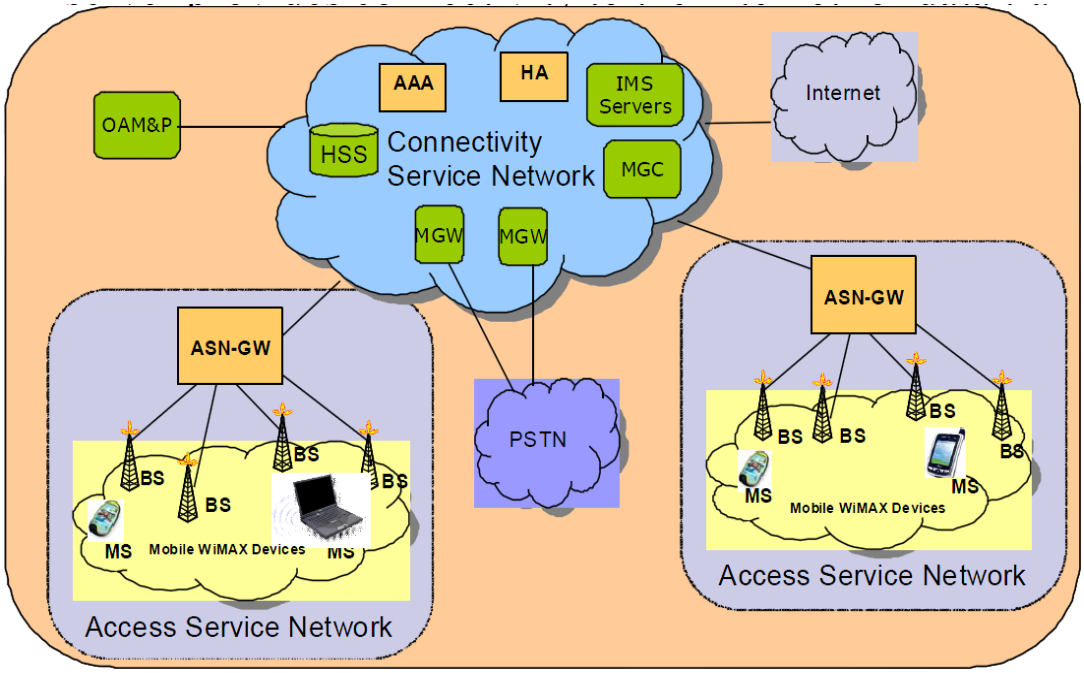

Figure 11. Mobile WiMAX network, from [9].

WiMax [9] is an all-IP, flat network whose basic elements are the Access Service Network (ASN) and the Connectivity Service Network (CSN). From the following figure that depicts the Network Reference Model, it is obvious that there may be two CSNs, one in the home Network Service Provider (NSP) and one in the visited NSP, as well as multiple ASNs.

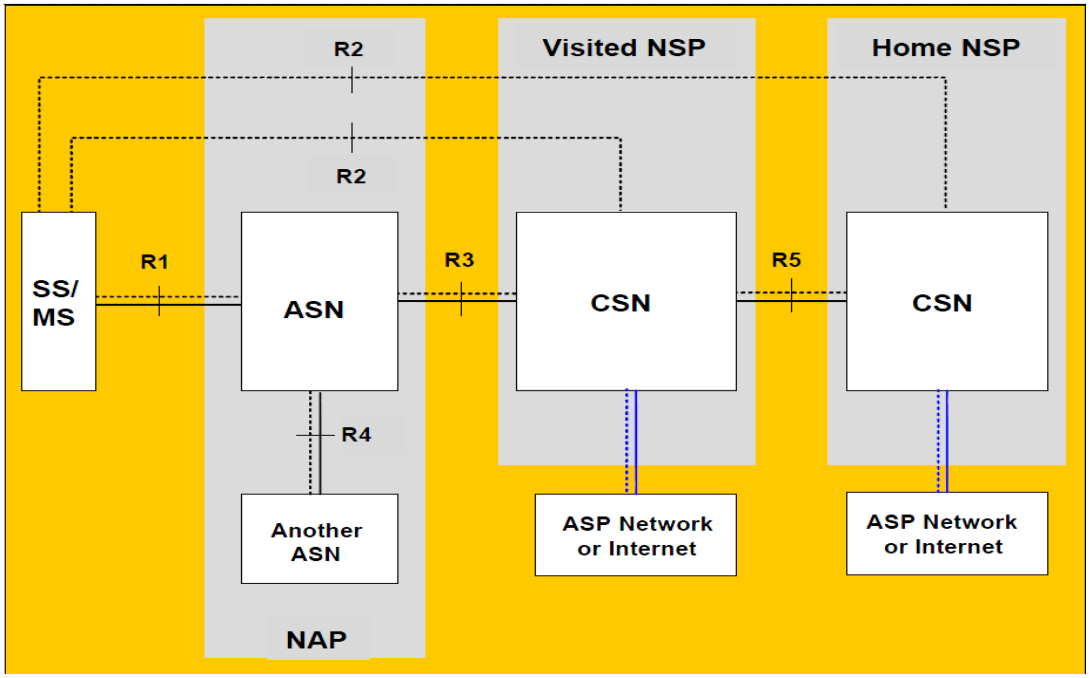

Figure 12. WiMAX network reference model, from [13].

Analyzing the main entities, the mobile stations are the entities through which the subscribers attain access to the network.

The Network Access Provider (NAP) consists of one or more ASNs. Every Access Service Network (ASN) forms the access service network comprised of the BSs and the ASN gateways that are connected over an IP infrastructure; it provides the set of functions that are related to access services. The ASN-GW enforces security, since MS user traffic is tunneled as payload between the BS and itself. Moreover the ASN-GWs offer message forwarding and mobility of MS.

The CSN may reside either on the home NSP or the visited NSP and it represents the network functions necessary for IP connectivity. The visited NSP CSN is the one that services the subscriber while the home NSP CSN is where the subscriber actually belongs. When the MS is not roaming there is only one NSP, the home NSP.

The CSN consists of many network elements, as depicted in Figure 11. First of all, the AAA server includes an AAA database where the mobile station profiles are stored. The server is responsible for authenticating the MS through messages that arrive from the ASN-GW. After authentication takes place, the mobile station's profile, along with QoS parameters, are sent to the ASN-GW. The Home Agent (HA), which provides global mobility and data transport to the Internet, processes the control signals from the ASN-GW, assigns a mobile IP address to the mobile station and keeps track of the IP payload. IP Multimedia System (IMS) servers are entrusted to process Voice Over IP (VOIP) calls inside the WiMAX network. Media Gateway Controllers (MGW) are used to provide access to the Public Switched Telephone Network (PSTN) if the call is terminated outside the WiMax network.

Lastly, in the case of multiple ASNs in an NAP, mobility is handled through the ASN-GWs. When a MS moves from a home NSP to a visited NSP, the AAA server takes care of the transfer of credentials and profiles from the home NSP to the visited NSP. The calls can be transferred in the same way when an MS moves from one BS to another served by the same ASN-GW [9].

b. WiMAX Security Architecture

The WiMAX security architecture [9, 14, 15] is based on 802.16 standards and addresses some known existing security problems in 3G networks. The architecture is based on an AAA framework that provides device, user, and mutual authentication between the MS and the NSP; global roaming; QoS policy control; and secure mobility management [14].

One of the basic security concepts in WiMAX is the introduction of a security sub-layer in the MAC layer for the wireless link between the BS and the MS. The IEEE 802.16e defines the standards for the security sub-layer and the security protocol stack, as depicted in Figure 13.

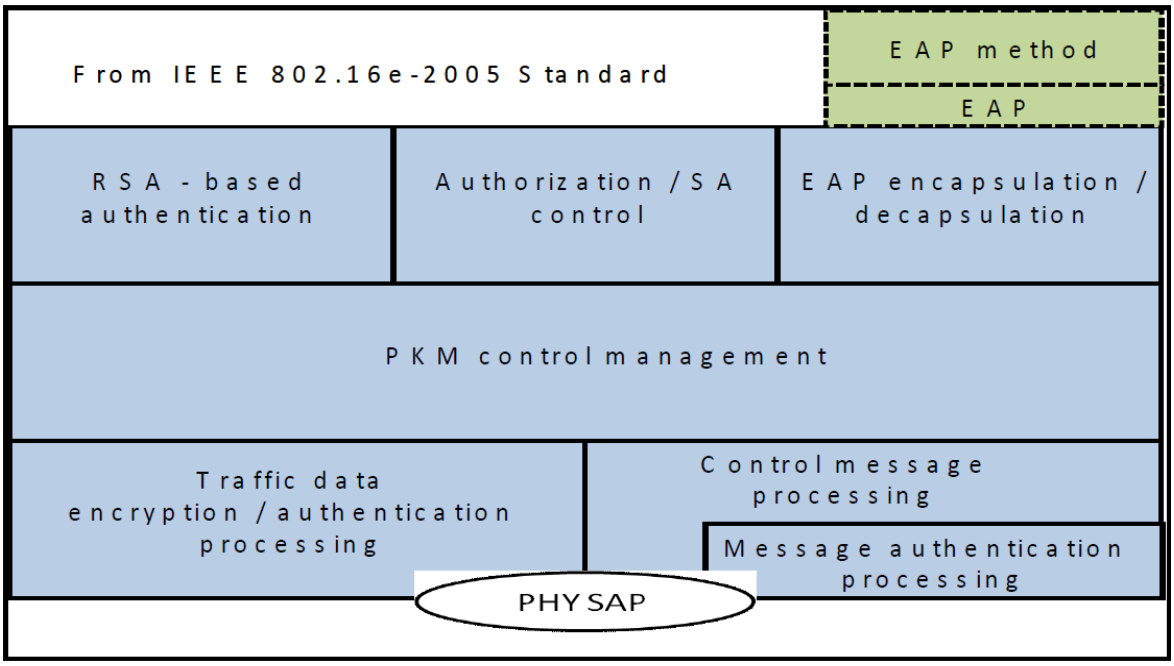

Figure 13. Security protocol stack for WiMAX 802.16e, from [9].

The security sub-layer is responsible for authentication and authorization, key management and distribution, and data encryption.

(1) Authentication. User and device authentication between an MS and the home CSN relies on Privacy and Key Management version 2 (PKMv2). There are two potential types of authentication [9, 14, 15]:

- RSA Based Authentication: WiMAX devices presuppose that credentials necessary for the authentication procedure (X.509 digital certificate) are loaded before the first time they are used and that these credentials are also programmed in the AAA server that resides in the home CSN. The X.509 certificate is issued by the MS manufacturer and contains the MS's public key and MAC address. The certificate is sent to the servicing BS during an Authentication Key (AK) request; it is validated by the BS and then the AK is encrypted using the MS's PK and sent back to the MS.
- Extensible Authentication Protocol (EAP) based authentication: the MS is authenticated by either a unique operator-issued credential (like SIM, USIM, user id and password) or the X.509 certificate. There are three EAP schemes: EAP-AKA for SIM based authentication, EAP-Transport Layer Security (EAP-TLS) for X.509 based

authentication, and EAP-Tunneled Transport Layer Security with Microsoft Challenge-Handshake Authentication Protocol version 2 (EAP-TTLS MS-CHAP v2) to provide secure connections in roaming cases and also user credentials protection.

In Figure 14, the layering of the PKMv2 user authentication protocols is depicted.

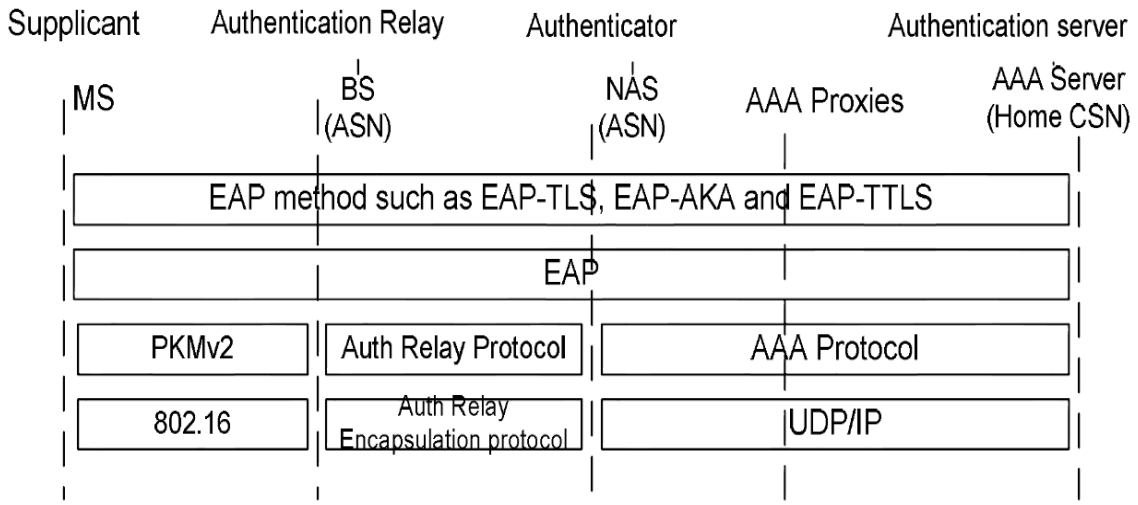

Figure 14. PKMv2 user authentication protocols, from [15].

PKMv2 transfers the EAP between the MS and the BS in the ASN. If the authenticator does not reside in the BS then the EAP is forwarded to the authenticator over an authenticator relay protocol. In the authenticator, the EAP messages are encapsulated in AAA protocol packets and sent to the AAA server in the home CSN through AAA proxies.

(2) User Authorization. The MS sends an authorization request to the BS, which includes an Authorization Key request and a Security Association IDentity (SAID) request, by sending the X.509 certificate encryption and cryptographic algorithms to the BS. The BS, after successful validation with the AAA server, responds by sending the AK, encrypted with MS's public key, along with a lifetime key and an SAID. The authorization by the AAA server

happens only the first time the MS associates with the BS. After that, the BS authorizes the MS without interacting with the AAA server [15].

(3) Key management/distribution and traffic encryption. The PKMv2 procedures [15] on which the WiMAX security relies are depicted in Figure 15.

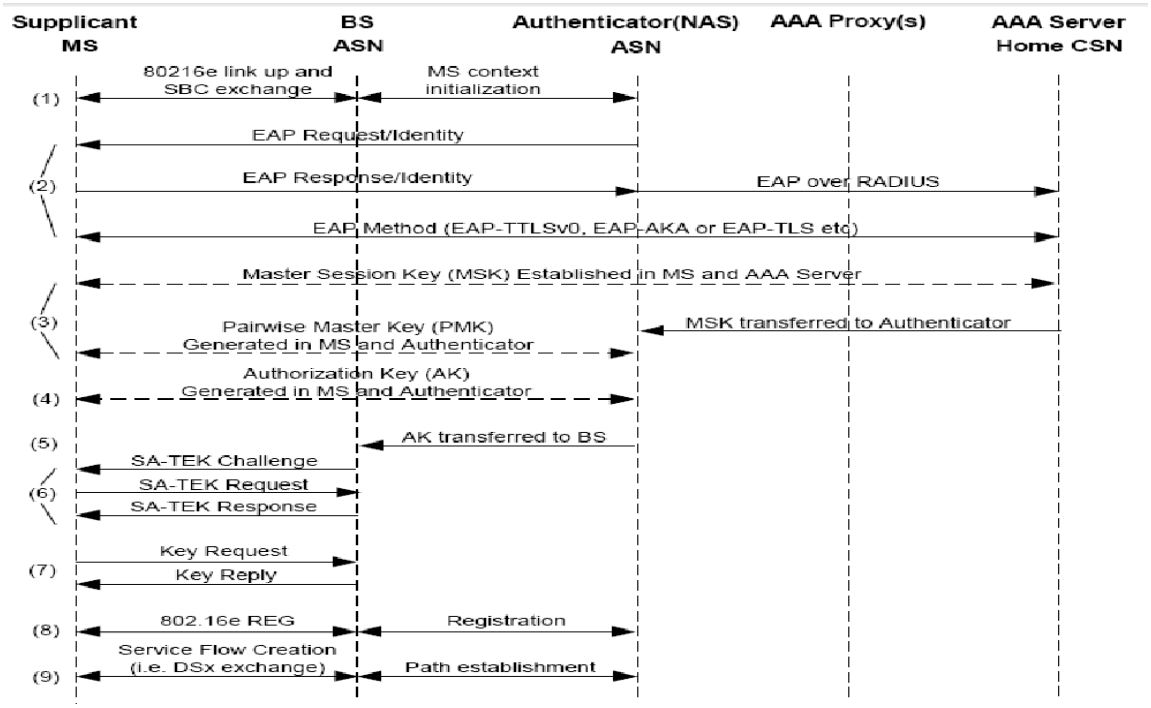

Figure 15. PKMv2 procedure during initial network entry, from [15].

In the first step, after successful ranging takes place, the MS and the ASN negotiate the session link capabilities by exchanging Subscriber Station Basic Capabilities (SBC) messages, including the PKMv2 security capabilities, authorization policy and device requirements. After the link is set up, the MS sends an initialization message to the Authenticator to begin the EAP sequence.

In the second step, the authenticator sends an EAP Identity/Request message to the MS. The MS replies by sending EAP response messages, which are forwarded through AAA proxies, to the AAA server. The

procedure is finished after receipt of one or more messages from the AAA server informing the user whether or not the authentication was successful.

In the third step, a Master Session Key (MSK) and an Enhanced Master Session Key (EMSK) (64 bytes or longer) are established between the MS and the AAA server. The MSK is also forwarded to the Authenticator by the AAA server and is used to generate the Pairwise Master Key (PMK) between the MS and the Authenticator. Further, the MS and the AAA server use the EMSK to generate mobile keys.

In the fourth step, a 160-bit Authorization Key (AK) based on the PMK is generated by both the MS and the Authenticator.

In the fifth step, the Authenticator transfers the AK and its context to the BS, which caches the information necessary for future action.

In the sixth step, a Security Association (SA) three-way handshake procedure takes place between the MS and the BS. The SAs are the security information that the BS and one or more MSs share to enforce security in their communication. The Traffic Encryption Key (TEK) that is used in the three-way handshake is a random number that is generated in the BS based on the AK. The 128-bit Key Encryption Key (KEK) is used to encrypt the TEK before the key transfer from the BS and it is generated based on the AK. The SA may belong to one of the following categories: primary, static, or dynamic. The MS establishes the primary SA during the initialization phase. The static SAs are established within the BS. Dynamic SAs are created and destroyed depending on the service flows. MS and BS compute Hashed Message Authentication Code (HMAC) values and compare them to those that are sent with the encrypted data in order to find integrity failures. The three-way handshake starts by the BS sending a Security Association Traffic Encryption Key (SA-TEK) challenge to the MS. The SA-TEK challenge includes the AK that is going to be used and the unique challenge value. After successful verification of HMAC, the MS sends an SA-TEK Request to the BS asking to be authorized to access potential SAs. The

three-way handshake finishes with the SA-TEK Response, where the BS identifies the primary and static SAs to which the MS is allowed access.

In the seventh step, the MS requests Traffic Encryption Keys, which are generated and encrypted using KEK, from the BS and sent back to the MS according to the procedure stated in the previous paragraph.

In the eighth and ninth step, the TEK registration takes place between the BS and the Authenticator, as well as service flows mapped onto an SA, completing the PKMv2 procedure [15].

THIS PAGE INTENTIONALLY LEFT BLANK

III. SECURITY ISSUES IN 3G, LTE AND WIMAX AND PROPOSED SOLUTIONS

In Chapter II, the necessary background regarding the security architectures of the three wireless networking technologies of interest was presented. This chapter begins by providing a security threats and vulnerabilities report. It then presents a few solutions that try to eliminate the weaknesses of these technologies.

A. SECURITY THREATS AND VULNERABILITIES

As technology improves, scientists try to eliminate known security threats and vulnerabilities. However, even if all known security issues are fixed, there may still be many issues that may not be well-known that make the technologies vulnerable to intruders. One factor is that every new technology is based on the previous one and interworking between the two technologies must be provided. This restriction makes it more difficult to address the security issues. In this section the security weaknesses, as well as the potential attacks, of the three wireless technologies of our interest are discussed.

1. UMTS Security Issues

The 3G security architecture was built on the concepts of GSM, but it addressed many weaknesses of GSM. Despite the security efforts that have been made to protect the 3G networks, there are some cases where the adversary can find opportunities to attack the networks.

a. Subscriber Identity Catching

The Network Access Domain is the most vulnerable part of 3G networks since it is responsible for protecting the wireless link. The wireless link is the most difficult to protect and most easy to intercept. One of the most common vulnerabilities of Access Security is catching the subscriber identity. In order to understand this kind of vulnerability, the cases where the IMSI is

provided are restated. Whenever the subscriber turns on his/her UE and whenever the correspondence between TMSI and IMSI (which is tracked in the VLR) is lost, the UE has to send the IMSI in the clear over the radio link. Thus, in the initial connection request message and when the VLR database loses synchronization with respect to the UE, the IMSI is sent unencrypted. A very simple mechanism that can be used by an attacker is depicted in Figure 16 [14]. The attacker impersonates a fake base station. During the start of the connection process, the victim uses the TMSI. After the user fails to be recognized, the fake VLR asks the user to identify himself/herself and the victim sends the actual IMSI in the clear. Then the attacker disconnects, having acquired the subscriber's IMSI [16], [17].

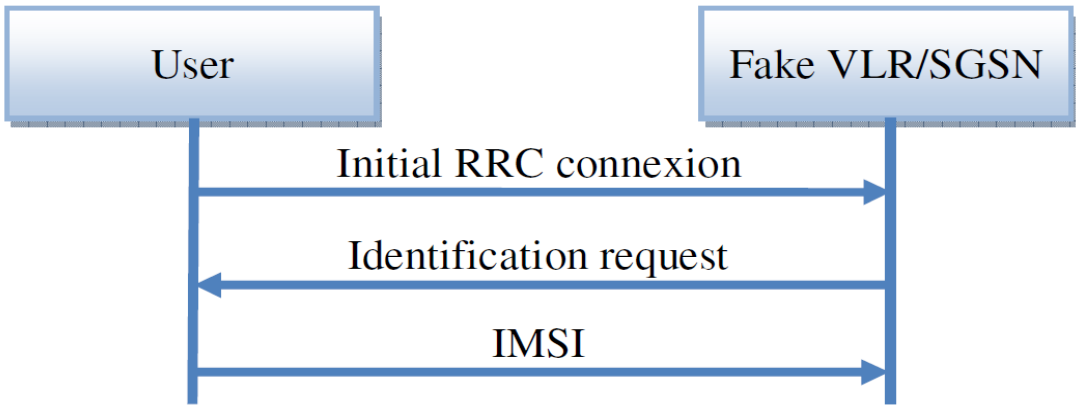

Figure 16. Obtaining IMSI, from [16].

b. Secret Key and Confidentiality Key and Integrity Key Exposure

The secret key, K, as well as the confidentiality and integrity keys, CK and IK, respectively, are vulnerable and can be disclosed to an adversary through cryptographic attacks. An adversary can use some data that is transmitted on the radio access link and gain access to derive the K, CK, and IK, as depicted in Figure 17. If the attacker wants to obtain the CK and IK he will use

the protected information that is transferred between the user and the network and apply a cipher-text-only attack to the encryption function, f8, and integrity function, f9 [18].

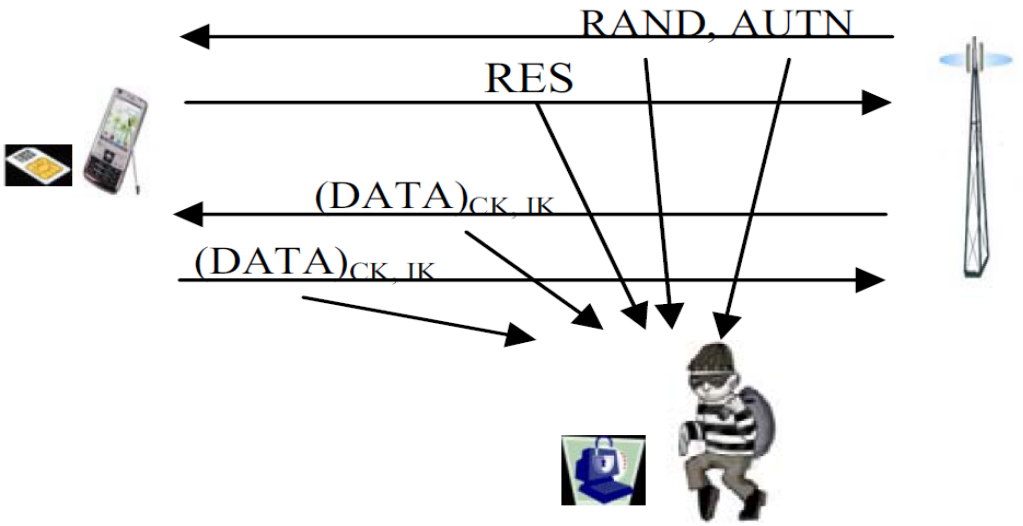

Figure 17. Attacks on the radio link, from [18].

When the attacker wants to derive the secret key, K, he uses the messages that are exchanged during the AKA procedure and applies cryptographic attacks to the security functions, f1 and f2. In Figure 18, the data that are exposed are the RAND, AMF, MAC, and XRES, as well as the security functions, f1 and f2. Thus, the adversary can mount a cipher-text-only attack against f1 and known plaintext attack against f2 [18].

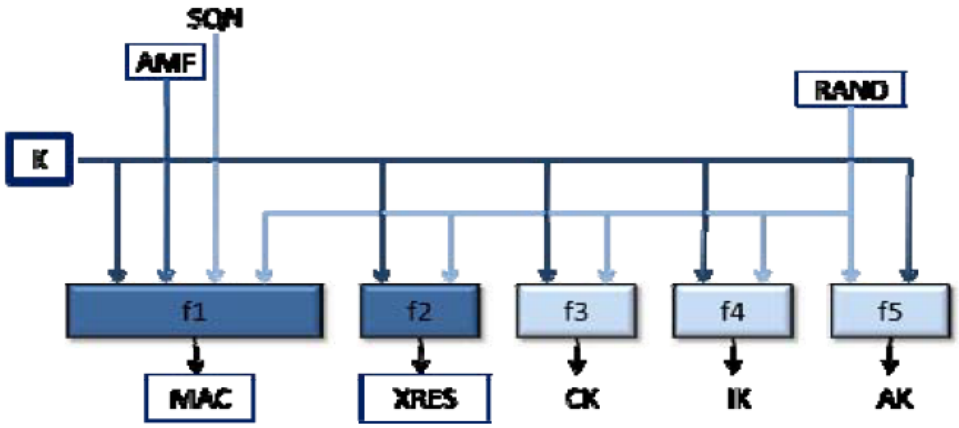

Figure 18. Exposure of security functions to cryptographic attacks, from [18].

c. User Specific DoS by Modifying Initial Security Capabilities of ME or Authentication Parameters

This denial-of-service (DoS) attack presupposes that the IMSI is revealed to the attacker using Subscriber Identity Catching discussed earlier. Thereafter, when the specific user associated with that IMSI makes a connection request the adversary modifies the security capabilities of the ME. That modification, which is not integrity protected, remains undetected until the exchange of security capabilities takes place and the connection procedure terminates. This procedure may consume bandwidth and resources making effective a user DoS attack. Using the same technique, the attacker can modify the un-encrypted, non-integrity protected authentication parameters, AUTH, RAND or RES, preventing the authentication of the network and the user. Thus, another DoS attack is created [17].

d. DoS Using Connection Reject Message

This DoS attack also presupposes that the IMSI is revealed to the attacker using Subscriber Identity Catching. When the user tries to connect to the network he makes use of one of his unique identifiers, TMSI, P-TMSI, or IMEI. The attacker then rejects the connection request. The UE compares the initial UE identity with the one received in the reject message in order to confirm the

legitimacy of the message. If the identities are the same, the connection procedure is terminated. Otherwise he ignores the message. Thus, a DoS attack against the user becomes feasible [17].

e. DoS by Flooding the HLR/AuC

As described in the previous section, the attacker utilizes Subscriber Identity Catching and builds an IMSI database. Then he uses an automatic procedure to generate a connection request per IMSI. The rogue VLR sends all the IMSIs included in its database except the one that is already connected to the HLR/AuC. The HLR validates the IMSIs and computes the five Authorization Vectors (AVs) per IMSI. This is a time-consuming process, especially if the number of the IMSIs is large. The AVs are forwarded to the VLR which selects one AV per IMSI and sends the RAND and AUTN for authentication. The attacker, of course, will not be authenticated since he does not have the key needed to compute the RES, but he has already exhausted the computing resources and bandwidth of HLR/AuC by flooding a significant number of connection requests. Thus, a DoS to new users becomes feasible [17].

f. Redirection Attack

For the redirection attack [19, 20], it is assumed that an adversary is operating a device that has base station capabilities as well as mobile station emulating capabilities. Thus, the attacker can impersonate both a base station and a mobile station at the same time. Therefore, a mobile station can connect to the false base station and the false base station can connect to a legitimate base station. When a user is in the area of his home network and tries to connect to a genuine base station, the attacker intercepts the connection and enables a connection with the false base station. Thereafter, it sends a connection request to a legitimate serving network on behalf of a victim’s mobile station and transfers securely the messages between the victim and the serving network (SN). This happens since the authentication between the victim and the serving network is

feasible, and the communication is protected through established keys. The redirection attack [20] is depicted in Figure 19.

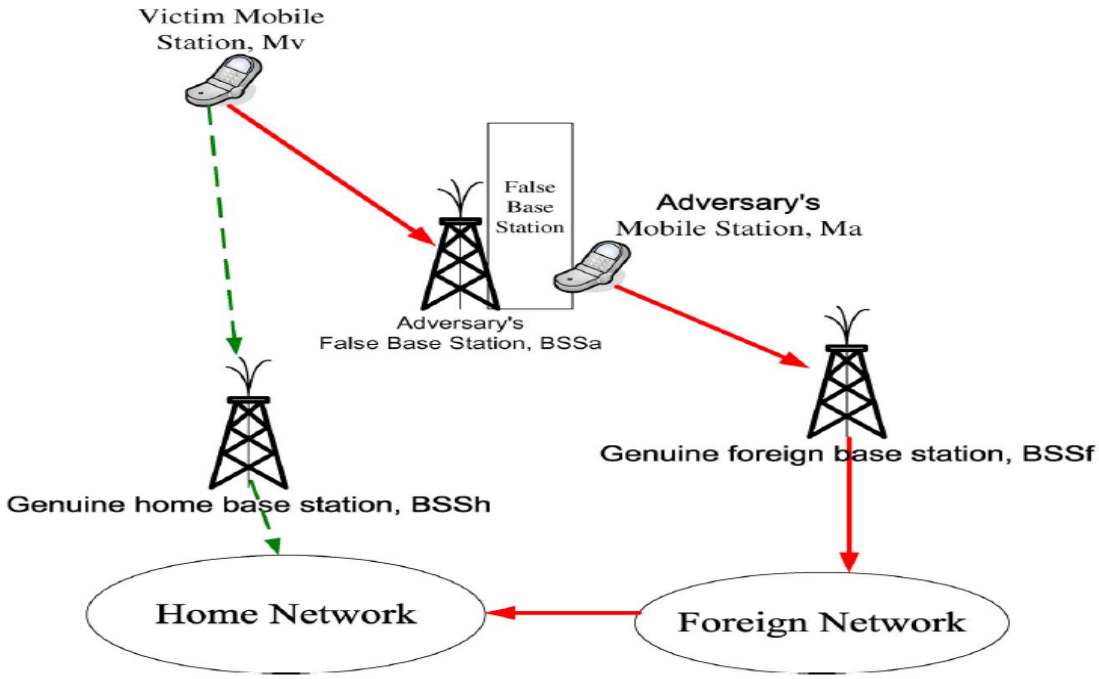

Figure 19. Redirection attack in UMTS AKA, from [20].

The redirection attack causes billing issues to the user since he actually uses the services of the home network but pays for roaming through the foreign network. Neither the home network nor the victim can identify the redirection attack. Moreover, the intruder can forward the traffic through a network that has no or weak encryption, enabling the adversary to eavesdrop on traffic [19].

g. Man-in-the-Middle Attack

In the 3GPP-AKA, the SN is not authenticated during the authentication process. Thus, the MS or the HN cannot determine the legitimacy of the SN. An attacker, who can mount a Man-in-the-Middle attack between the MS and the SN over the wireless network, can exploit that limitation. In this way, the attacker is interjected between the legitimate entities and can have access to

the authentication messages as well as the traffic messages, which are then subject to alteration. Thus, DoS attacks are possible. A few attack models over the wireless network were developed and analyzed in depth in [21].

In one of the models of wireless network attack, the attacker impersonates the SN and sends a reject message to the MS during the authentication procedure, enabling a DoS attack since the MS believes that the SN is legitimate. In another model of wireless network attack, the attacker modifies the RAND that it receives from the SN and forwards a false F-RAND to the MS. When the MS realizes that the SN is fraud, it terminates the authentication procedure and the DoS attack succeeds. In a third model, the attacker modifies the RES that is received from the MS and sends a false F-RES to the SN. The SN identifies that the MS is fraudulent, terminates the authentication process, and the DoS succeeds. The models are depicted in the following figures:

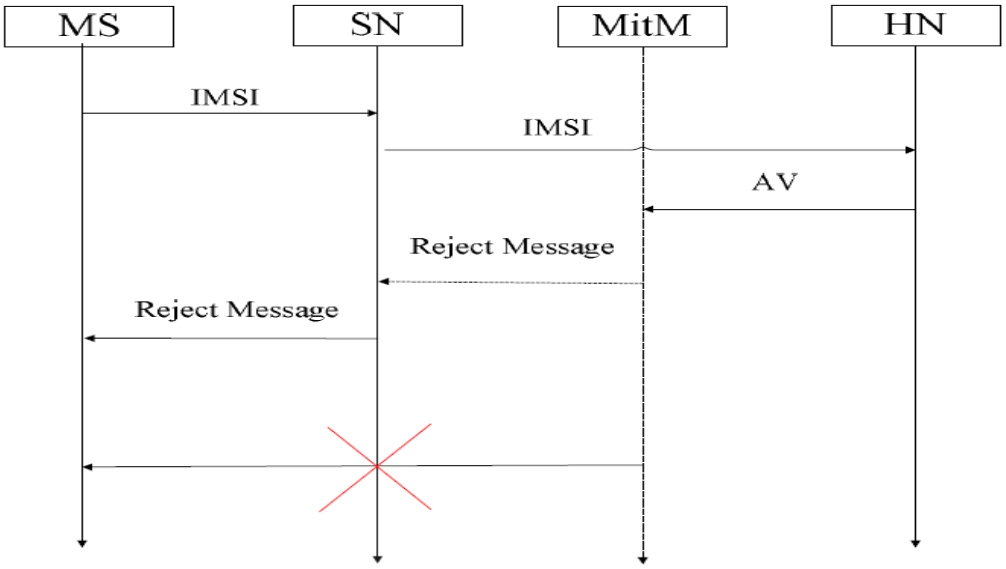

Figure 20. First attack model using authentication rejected message (ARM) in wireless network, from [21].

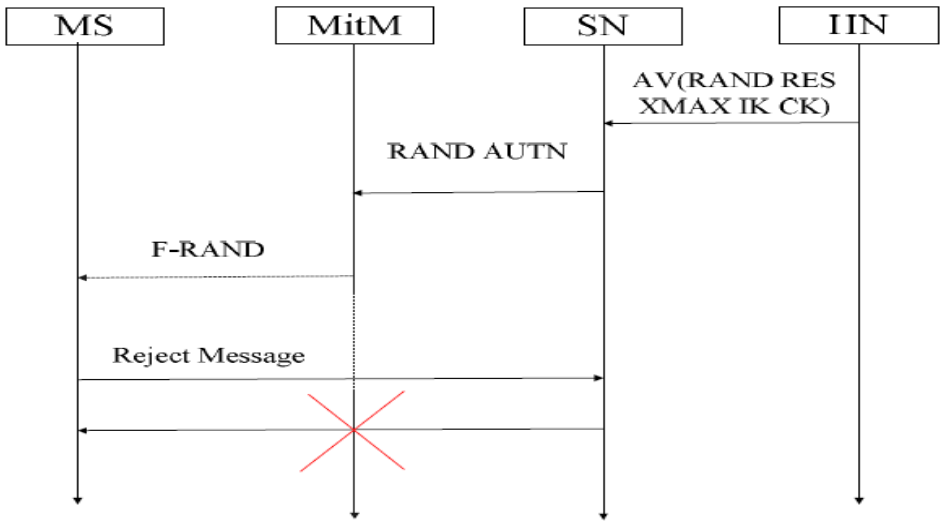

Figure 21. Second attack model using RAND modification in wireless network, from [21].

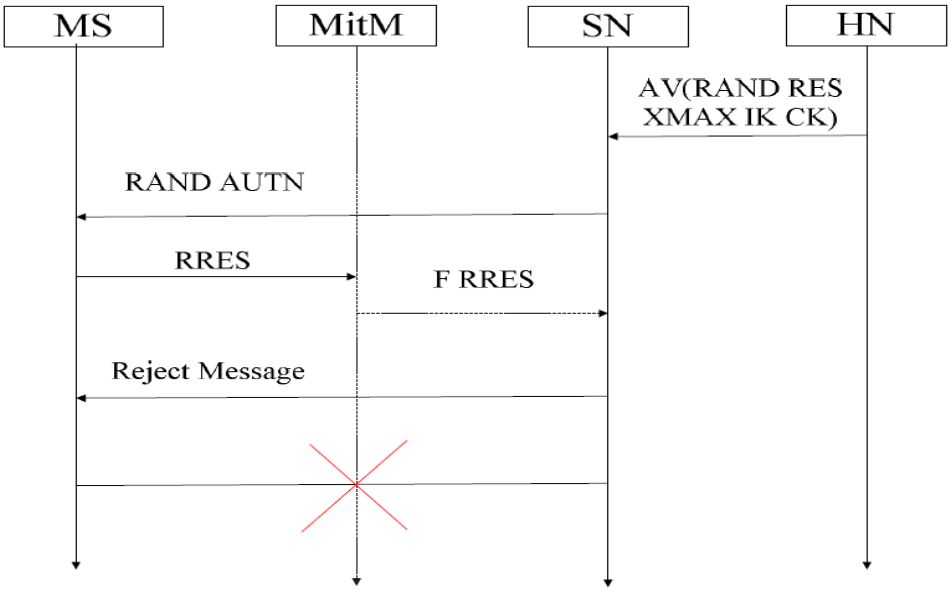

Figure 22. Third attack model using RES modification in wireless network, from [21].

h. Man-in-the-Middle Attack and Base Station Impersonation of Combined UMTS/GSM User Equipment

In this case, the attacker takes benefit of the fact that GSM does not offer integrity protection. Thus, during UMTS - GSM interworking the attacker can force the subscriber to use no encryption. The attack assumes that the adversary already knows the victim's IMSI and security capabilities using one of the methods described earlier in this chapter.

At first the attacker impersonates the victim MS and sends to the visited network the security capabilities as well as the TMSI during connection set-up. If the TMSI is not recognized the MS responds to the identity request by sending the victim's IMSI. Following this, an authentication request between the visited and the home network takes place and, after successful authentication, the RAND and AUTN are forwarded to the adversary, who finally disconnects. This procedure is depicted in the Figure 23:

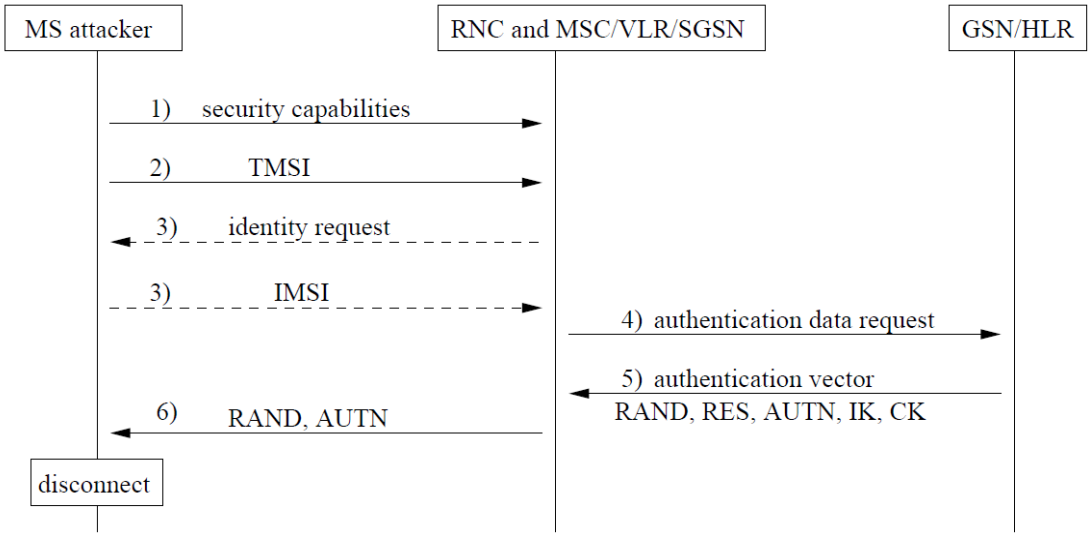

Figure 23. Attacker obtains currently valid AUTN, from [25].

Subsequently, the adversary impersonates a valid GSM base station. The attacker lures the target MS by sending its beacons with higher transmitting power causing a hand-off and establishes a connection with the MS.

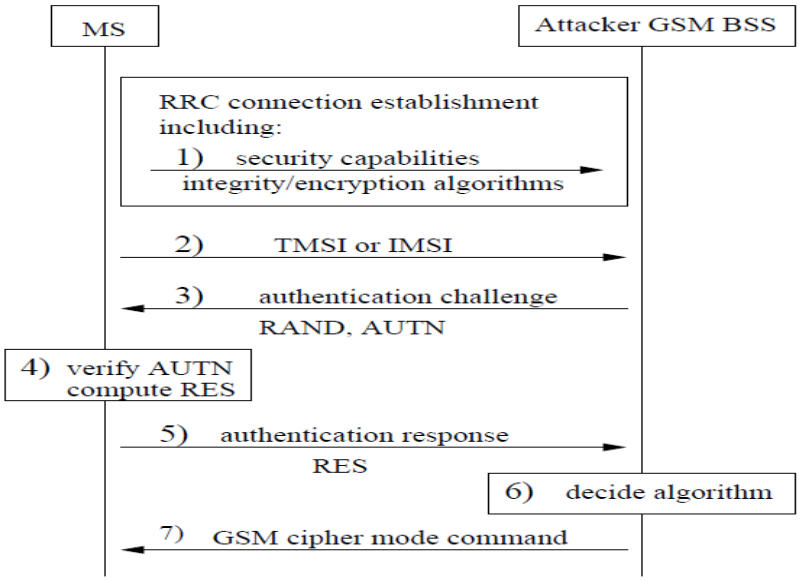

Figure 24. Attacker impersonates valid GSM base station, from [25].

During the connection establishment, the victim MS sends its security capabilities and its identifier, either the TMSI or the IMSI, to the attacker. The impersonated BS sends the RAND and AUTN derived from the real network and the victim computes the RES and replies, stating that the authentication token is verified. Thereafter, the false base station dictates the use of no or weak encryption to the MS. The MS accepts it, considering it is connected to a GSM base station. This procedure is depicted in Figure 24. Thus, the intruder succeeds in fooling the MS into not using encryption and can thereafter eavesdrop on all communication from that MS [25].

2. LTE Security Issues

The LTE security architecture and 4G technologies, in general, have significant differences as compared to the previous technologies; the most important of which is that they operate entirely based on TCP/IP architecture and protocols. Even if efforts were made during design to reduce the security issues caused by the open nature and IP-based infrastructure, there are still a few security issues with which to be concerned.

a. IP-based Vulnerabilities

LTE, since it is a flat, all-IP-based architecture, is vulnerable to well-known attacks associated with Internet. Thus, there is a danger of IP address spoofing, phishing attacks, DoS attacks, viruses, worms, spam mails and calls [10].

b. Base Station Attack

The fact that the Mobile Management Entity (MME) in LTE handles many eNBs and HeNBs (in LTE-A), along with the flat IP-based architecture, makes it easier for an adversary to attack a base station than in UMTS. In the UMTS the serving network only handles a couple of Radio Network Controls (RNC) in a hierarchical way, while in LTE a direct path is offered to attackers due to its being an all-IP network. Moreover, if a base station is compromised, the whole network will be in danger because of the nature of IP-based architecture as it is explained in [22, 23, 24]. Furthermore, with the introduction of HeNB in LTE Advanced networks, a few more threats arise. The HeNB is a small, low-power cellular base station that has the eNB's functionality. It is typically intended for use in residences or small businesses, offering increased indoor coverage for voice and high-speed data service. However, HeNB may use an insecure link to connect to the SGW, offering an attacker vector to adversaries. For example, once an HeNB is compromised, the attacker can create a fraudulent version with dual functionality that can impersonate a base station and a user at the same time [10].

c. HeNB Weakness

Beyond the weaknesses of HeNB associated with IP-based attacks, it is vulnerable to a few other threats. Since there is no mutual authentication between the UE and the HeNB, there is a risk of eavesdropping, masquerading and Man-in-the-Middle attacks. Moreover, since it is exposed to the public Internet, DoS attacks become easier to make against a HeNB. Lastly, HeNB may be open to physical intrusions because of its placement in unsecure areas [10].

d. Handover Authentication Vulnerabilities

A few concerns arise regarding the handover authentication procedure due to the multiplicity of mobility scenarios between eNBs and HeNBS. Even if the 3GPP committee defined a few mobility and authentication scenarios [10] problems still may arise in the case that the base stations are handled by different MMEs. Furthermore, more network security threats arise due to the heterogeneous networks interoperating with LTE. This is a problem especially during the process of transferring an ongoing call or data session from one connected cell of the core network to another, called handover or handoff. Even if the 3GPP committee has made suggestions for a secure handover between the E-UTRAN and non-3GPP access networks during roaming [10], the UE has to go through the whole access authentication procedure with the new network before it completes the handovers. Thus, many messages have to be exchanged with the AAA server, causing more handover delay. Furthermore, the key derivation procedures that were analyzed in [10] concluded that multiple key management mechanisms are required, which increases the overall complexity. While the potential delays may impact user satisfaction, the added complexity of key management and multiplicity authentication actions expose issues that exploited by adversaries to consume network resources of the core network or other access networks [10].

e. MME Buffer Exhaust - HSS Computational Power Exhaust

During the authentication procedure, as depicted in the Figure 10 in Chapter II, the MME at first has to forward the messages from the UE to the HSS even before the UE is authenticated by the MME. Then after receiving the authentication data response from HSS, the MME sends the user authentication request to the UE. After receiving the user authentication response, the MME can authenticate the UE. The attacker might exploit this process by impersonating a legitimate user and sending fake IMSIs to cause DoS attacks to the HSS and the MME, causing the HSS to compute too many authentication vectors and the MME to exhaust memory buffer allocations waiting for long periods of time for the UE to send the user authentication response [10].

f. IMSI Catching

In the EPS-AKA protocol, the authentication starts with the user identification in which the UE sends the IMSI in plain text. Moreover, if there are synchronization failures when the UE roams to a new MME or the current MME cannot be contacted, a user identity request message is sent from the MME to the UE, to which the user has to send his IMSI, also in the clear, if no other temporary identifier is valid. The adversaries can then collect the users' IMSIs and use them to mount potential attacks, such as DoS [26].

g. User Equipment Tracking

There are a few vulnerabilities that lead to user tracking. These vulnerabilities are categorized as follows:

(1) Tracking User Temporary ID. The attacker may track the user's temporary ID and then link the temporary ID to the user. The adversary may link to the user when the user tries to connect to a compromised server and inserts his user name. Another way is to wait for the user to transmit his permanent or temporary ID when transmitted in plaintext [26].

(2) Exploiting the Linkability of IMSI/TMSI and RNTI. The adversary can track all the traffic that is associated with the temporary identifiers since temporary identifiers are transmitted over the air interface. Moreover, the attacker can impersonate an eNB and request the user's IMSI. After having derived the IMSI, the adversary can backtrack along the user's entire trace [26].

(3) Serving Network Authentication. The attacker can impersonate a serving network and simply forward the messages it receives from associating UEs. Thus, it can work as a fake serving network and after confidentiality protection is enabled it can intercept and decrypt encrypted data [26].

h. Wired Link Weakness

The wired link through which messages are transferred between the network entities is unprotected. All the messages are transferred unencrypted. Thus, there is a potential risk of an attacker intercepting and deriving the Authentication Vectors that are transferred from the HSS to the MME [27].

i. Symmetric Key Weakness

The authentication between HSS and UE is based on symmetric key encryption. Since the session cipher key can be disclosed to an attacker through the network, the resultant communication's security becomes questionable [27].

j. Service Network Identity (SNID) catching

The Service Network Identity is transmitted in plaintext over the air interface as well as on the wired link. Thus, it becomes possible for an attacker to derive the SNID and mount a false base station or non-legitimate network attack [27].

3. WiMAX Security

There are a few issues that lead to WiMAX insecurity. For example, the first steps of the procedure through which the MS initially accesses the BS are not secure. Some of the issues and potential attacks that arise in WiMAX security are provided below:

a. Unencrypted Management MAC Messages

Since management MAC messages are never encrypted, and sometimes not authenticated, there is a potential for an adversary eavesdropping the messages exchanged, deriving useful information, and mounting Man-in-the-Middle and DoS attacks. A few examples are provided in [9] and [28], as well as their impact.

In general, during initial network entry an MS exchanges information such as security settings, power and configuration settings, mobility parameters, and MS capabilities with the BS. Thus, an adversary can derive the information by just listening to the traffic. Since these messages are unauthenticated, unencrypted, and stateless, a Man-in-the-Middle attack becomes feasible and enables the attacker to derive this useful information. Moreover, almost all management messages are sent unencrypted, except the key transfer messages. Therefore, the attacker can eavesdrop on those messages, create an MS profile including the obtained information, monitor the MAC address, and associate the derived information to the user equipment [9], [28], [29], [30].

b. Unauthenticated Management Messages

Some management messages in IEEE 802.16e are integrity protected using either Hash Based Message Authentication Code (HMAC) or Cipher Based Message Authentication Code (CMAC). However, there are a few messages that are not authenticated at all, causing security vulnerability. In particular, the broadcast messages are difficult to protect since there is no

common key. Even if there was, it would be useless since the mobile stations sharing the key can easily counterfeit the messages. Thus, the communication between the BS and the MS can be intercepted [9], [28], [29], [30]. An extensive analysis regarding the unauthenticated management messages, specifically, is provided in [28], [30].

c. Interleaving Attack

In this case, the attacker impersonates an MS, sends an Authorization Information message and then an Authorization Request message that contains information which has been derived from previously intercepted sessions of the targeted MS. To these messages the attacker receives an Authorization Reply message, which is supposed to be verified by the MS by sending back an Acknowledgement response. However, the adversary, not having the MS's private key to decrypt the Authorization Reply message, cannot create the Acknowledge message. Instead, the adversary starts pretending to be the BS and communicates with the legitimate, targeted MS. By using the previously derived Authorization Reply message it solicits and receives the Authorization Acknowledgement message from the legitimate MS. Then the adversary uses the Authorization Acknowledgement message to complete the pending uncompleted session with the BS. In this way, the attacker authenticates himself with the BS and mounts a Man-in-the-Middle attack. Even if attacker has no access to the AK or TEK and so does not have the ability to decrypt or create encrypted messages, he can still act to forge or drop unprotected messages [28].

d. Authorization Vulnerabilities/Replay Attacks

The authentication/authorization protocol is vulnerable to replay attacks. If the Authentication Reply message is lost the Subscriber Station SS resends an Authentication Request message. The BS cannot determine if the SS did not receive the Reply message or if it is only a replay attack; thus it responds. An attacker can exploit this and flood the BS with Authentication Request messages limiting the BS‘s computing power. Moreover, the authorization

message does not include any digest so it cannot provide a means of integrity check and ensure the message was not modified [9].

e. Shared Keys in Multicast and Broadcast Service

The Multicast and Broadcast Service (MBS) offers the opportunity of distributing data to multiple MSs simultaneously by using one single message. The broadcast messages are encrypted symmetrically with a key shared among the group members, called the Group KEK (GKEK). This key is generated and sent from the BS to the group of MSs after the MSs successfully authenticate. Each MS can decrypt and encrypt messages using the GTEK. A rogue MS might pretend to be the legitimate BS. Moreover, since the GKEK is available to every group member, an adversary can use it to update the GTEK and distribute its own GTEK to the whole group through the Multi and Broadcast Rekeying Algorithm (MBRA), which is responsible for GTEK distribution. With this an adversary can fool the rest of the group members, which accept the new key and are no longer able to encrypt the traffic coming from the legitimate BS. Another issue that is related to the GTEK is that the member, after receiving the GTEK, can decrypt any traffic sent in the past from the moment when the GTEK first became active, assuming of course the rogue adversary had eavesdropped the traffic, as well as any additional traffic sent until GTEK lifetime expires [28], [29], [30].

B. PROPOSED SOLUTIONS

The system designers tried to eliminate the security weaknesses inherent in the three wireless technologies. There are many solutions that try to countermeasure one or more of the security threats. In the remainder of this chapter some of these solutions are discussed. These solutions try to mitigate the threats described in the first part of the chapter. Some of them are improved protocols of already existing solutions.

1. UMTS Security Enhancement

After 3GPP came out with the proposed security scheme for 3G, which was actually offering a relatively sufficient security, the academics tried to identify its threats and find ways to make it more secure. A few solutions addressing security threats and attacks on 3G are provided.

a. New Defense Strategy Model

Trying to defeat the Man-in-the-Middle attacks in the wireless networks, scientists created a few attack models and proposed a new security model of AKA in [21].

The security is based on the exchanged messages' encryption with public and private keys for the SN. The authentication between the MS and HN is based on a shared secret key, K_{hm}, with which the authentication messages between the HN and the MS are encrypted. First, the MS generates a random number, R_1, and determines the identity of VLR, ID_v. Next, it sends a message that includes its IMSI, R_1, and ID_v to the SN, which in turn forwards the message to the HN. The HN, after receiving and decrypting the message with the key K_{hm}, identifies the MS, generates a random number, R_2, the confidentiality and integrity keys, CK and IK, respectively, and the expected response, XRES, and sends them encrypted with the SN's public key to the SN. The HN sends the R_1, R_2, XRES and public key of the SN (K_{UV}) to the MS in a message encrypted by K_{hm}. Then, the SN decrypts with its private key the message received from the HN, derives the CK, IK, XRES and R_2, encrypts the R_2 with its private key and send it to the MS. The MS, after decrypting the messages received from the HN, derives the R_1 with which it identifies the HN; if the derived value of R_1 is incorrect, the MS stops the authentication. Otherwise, after obtaining the public key of the SN, it decrypts the message from the SN, which includes the nonce R_2 by which it identifies the SN. After decrypting the message received from the SN with K_{UV}, it compares the values of R_2 received from the HN and the SN to verify the SN identity. If they are different, the MS terminates the authentication.

Otherwise, it computes CK, IK, and RES and sends the RES to the SN where a comparison of RES and XRES takes place. If they are identical, authentication is completed, otherwise it fails [21]. The proposed model is depicted in Figure 25.

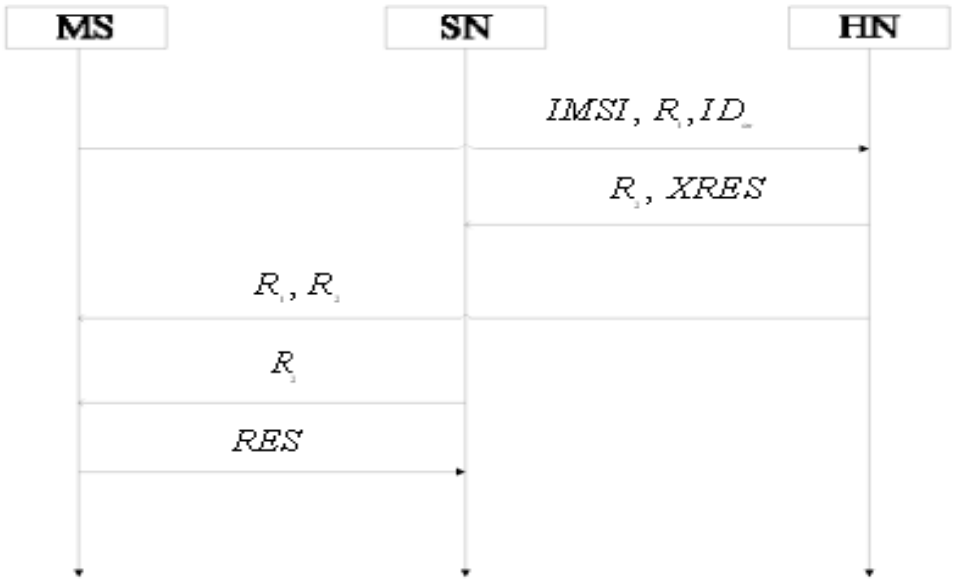

Figure 25. The proposed defensing model in 3GPP-AKA, from [21].

This model provides mechanisms for authenticating the SN and the HN in the wireless and the wired network through encrypting the messages and preventing Man-in-the-Middle attacks that were feasible on the Access Domain [21].

b. Enhanced EMSUCU Protocol

The solution, described in [18], was based on the Enhancement Mobile Security and User Confidentiality for UMTS (EMSUCU) solution that is described in detail in [31]. The proposed solution reflects the same philosophy as the one in [28] and it assures adequate protection for the secret key, K. The Enhanced EMSUCU mechanism [18] is depicted in Figure 26:

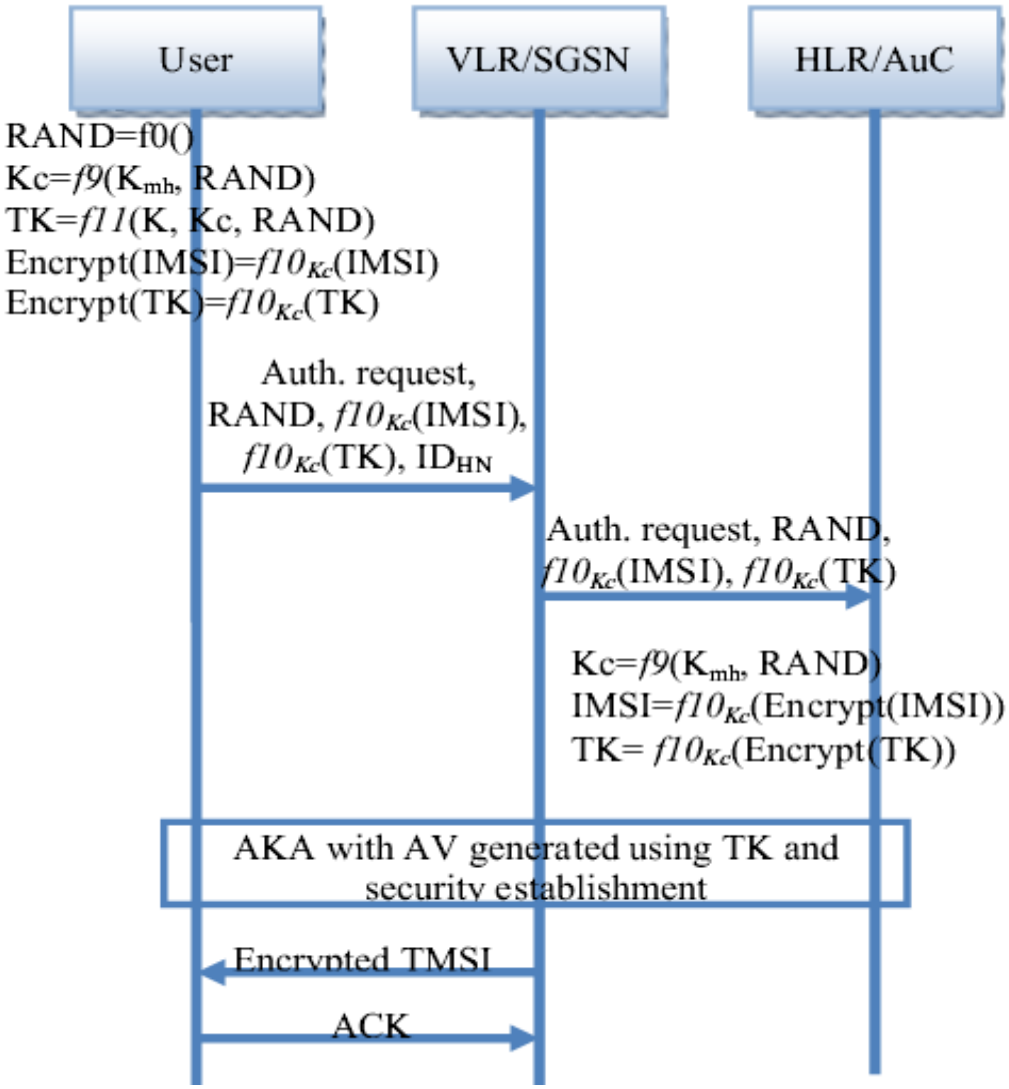

Figure 26. Enhanced EMSUCU, from [18].

The Enhanced EMSUCU proposes a solution that prevents IMSI disclosure over the air interface. The secret key, K, and the USIM card are used to encrypt the IMSI. The protocol assumes that there is a shared secret key, K_{mh}, for the home network that is known by the HLR/AuC and every USIM registered with it. A function, f0, is used to generate a random value, RAND, and an integrity one-way hash function, f9, to produce a cipher key, K_c, based on the RAND and K_{mh}. Another function, f11, is implemented on the USIM and generates a secret Temporary Key, TK, based on K, K_c and RAND every time EMSUCU is executed.

The TK is later sent encrypted to the HLR/AuC and used for AV generation. The one-way hash function, f10, along with Kc, is used to encrypt the mobile host's IMSI and TK. The key, $K_{mh,}$ and the functions, f9 and f10, are shared between the MS and the HN and are stored in the SIM card for the mobile user and in the HN for that user. The procedure starts by generating the RAND, K_c and TK, encrypting the IMSI and the TK and then sending an authentication request to VLR with these values and the identity of home network, ID_{HN}, so that the VLR knows to which HLR to forward the message. On the HLR side, after receiving these values, the HLR generates the K_c using the RAND, K_{mh} and f9 function. The K_c is then used to decrypt the IMSI and the TK. Then the AKA procedure takes place using the established TK. Thus, the security of the secret key, K, is increased since the adversary has to use a cryptographic attack and derive the key from the equation, which is not practically feasible.

$$f10_{Kc}(TK) = f10_{Kc}(f11(K, K_c, RAND)) = f10_{Kc}(f11(K, f9(K_{mh}, RAND), RAND))$$

Moreover, the Enhanced EMSUCU suggests that before the lifetimes of CK and IK expire a new AKA should be initiated. In this way, before the keys, CK and IK, are deleted, they will be used to protect the messages during the next AKA procedure. This process is simple to implement by just changing the order of the events in the UMTS protocol.

Lastly, the Enhanced EMSUCU suggests increasing the key size of K from 128 bits to a minimum length of 256 bits, which conforms to the NIST recommendations, increasing the key security [18].

c. S-AKA Protocol

The S-AKA protocol offers security against redirection and Man-in-the-Middle attacks and is described in detail in [20]. It is based on the UMTS AKA and assumes that the SGSN can handle user authentication securely, the link between the SGSN and the HLR/AuC is secure, and that each MS shares a secret key and cryptographic functions with its HLR/AuC. It consists of two phases [20]:

- In the first phase, called S-AKA-I, the SGSN tries to obtain an AV from the HLR so that it can authenticate the MS without interacting with the HLR during the second phase. In the first message, a new counter, called ACCm, which increases every successful authentication, is used to enforce freshness and derive the key, DK, from the SK. Further, the Location Area Identity (LAI), which is used to identify the BSS, the ACC_m and the derived DK are used for generating the MAC_m, the keyed authentication code enforcing message integrity protection. Thus, when the HLR receives the forwarded message from the SGSN it can verify that the LAI is the one recognized by the MS; otherwise, the HLR will reject the message. In the HLR, the ACC_m is compared with the HLR's counter, ACCh, and if the ACC_m is smaller than the ACCh it rejects the message, having identified the replay attack. If this test is passed, the HLR computes the DK and AUTN and sends it to the SGSN giving it the opportunity to authenticate the MS during future authentication procedures. The SGSN then computes the counter, ACCs, and the $AUTN_s$ and sends the last one to the MS, which verifies the SGSN, updates the ACC_m, and derives the CK, IK and XRES. If there is a GSM BSS involved, an extra key, PLK, is generated using a new function, f7, in both the MS and SGSN to provide confidentiality of the GSM BSS traffic and prevent Man-in-the-Middle attacks. Then the XRES is verified by the SGSN, a mutual authentication is enforced and the SGSN derives the keys CK, IK, and PLK, to enforce security between the MS and the SGSN. The first phase is depicted in Figure 26.

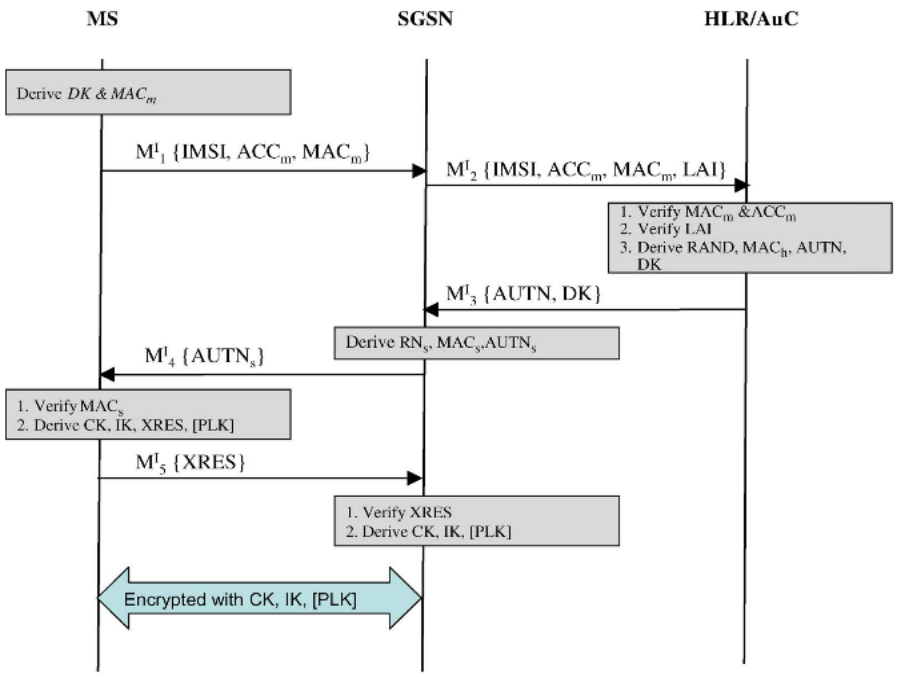

Figure 27. S-AKA-I. The SGSN obtains authentication vectors from HLR/AuC, from [20].

- In the second phase, called S-AKA-II, the procedure is similar, but there is no involvement of the HLR since the SGSN is authorized to authenticate the MS through AVs acquired from the first phase. The second phase is depicted in the Figure 28.

The procedures presented in this section, the functions that are used, and the procedures through which the counters ACCm, ACCs and ACCh are checked and updated are analyzed in detail in [20]. Lastly, the synchronized ACCm and ACCs can be beneficial for detecting potential DoS attacks initiated in the first message of the S-AKA second phase [20].

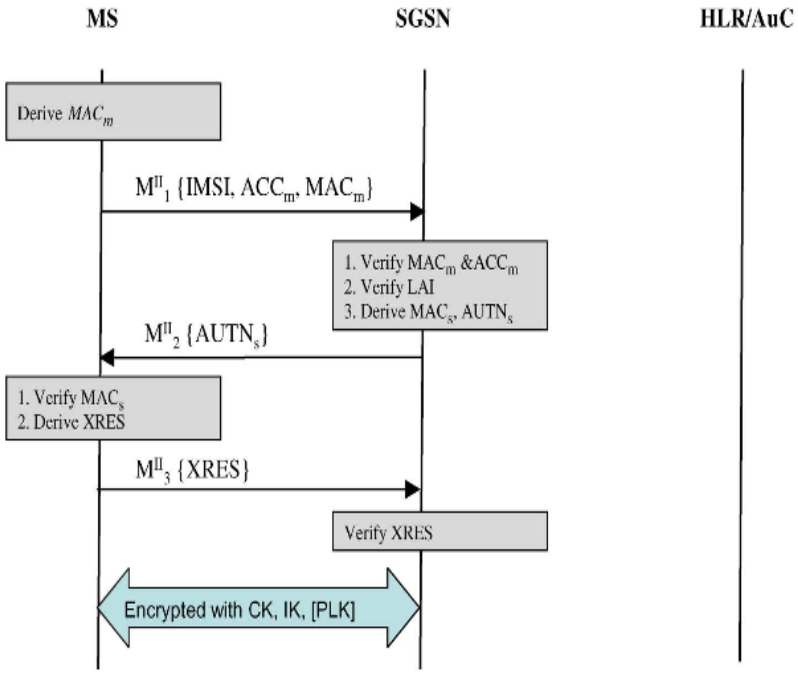

Figure 28. S-AKA-II. The SGSN mutually authenticates the MS, from [20].

2. LTE Security Enhancement

The LTE security architecture is IP-based. The various threats that became evident with the architecture forced the scientists to try to find new protocols to improve the state provided by the initial protocols. A few solutions are provided here.

a. Security Enhanced EPS-AKA

The Security Enhanced Evolved Packet System - Authentication and Key Agreement (SE EPS- AKA) [27] is an improvement of EPS-AKA, and it is based on the Wireless Public Key Infrastructure (WPKI). Before the communication between the UE, the MME, and the HSS starts, the digital certificate via CA is acquired through the procedure described in [32] and the public key are gained. The procedure of SE-EPS AKA is depicted in Figure 29.

First, the UE makes an access request and sends its IMSI, encrypted with the public key of the HSS, PK_H, stored in UE smart card, along

with the identity of the HSS, ID_{HSS}, to the MME. Then the MME uses the PK_H to encrypt its network identity, SNID, and sends it along with the message received from the UE to the HSS as an authentication request. The HSS decrypts the received message with its private key, SK_H, derives the IMSI and SNID, and checks if they are valid according to its database. If they are, the HSS generates an array of random numbers, RAND (1, 2,…,n), and a group of Authentication Vectors, AV(1,2,…,n), using the authentication algorithm depicted in Figure 30. Thus, using the authentication algorithm, the K_{ASME} and XRES are calculated and, thereafter, the AV via the equation [27]

$$AV = RAND||SNID||K_{ASME}||XRES$$

The HSS sends the AV along with the IMSI encrypted with MME'S public key back to the MME.

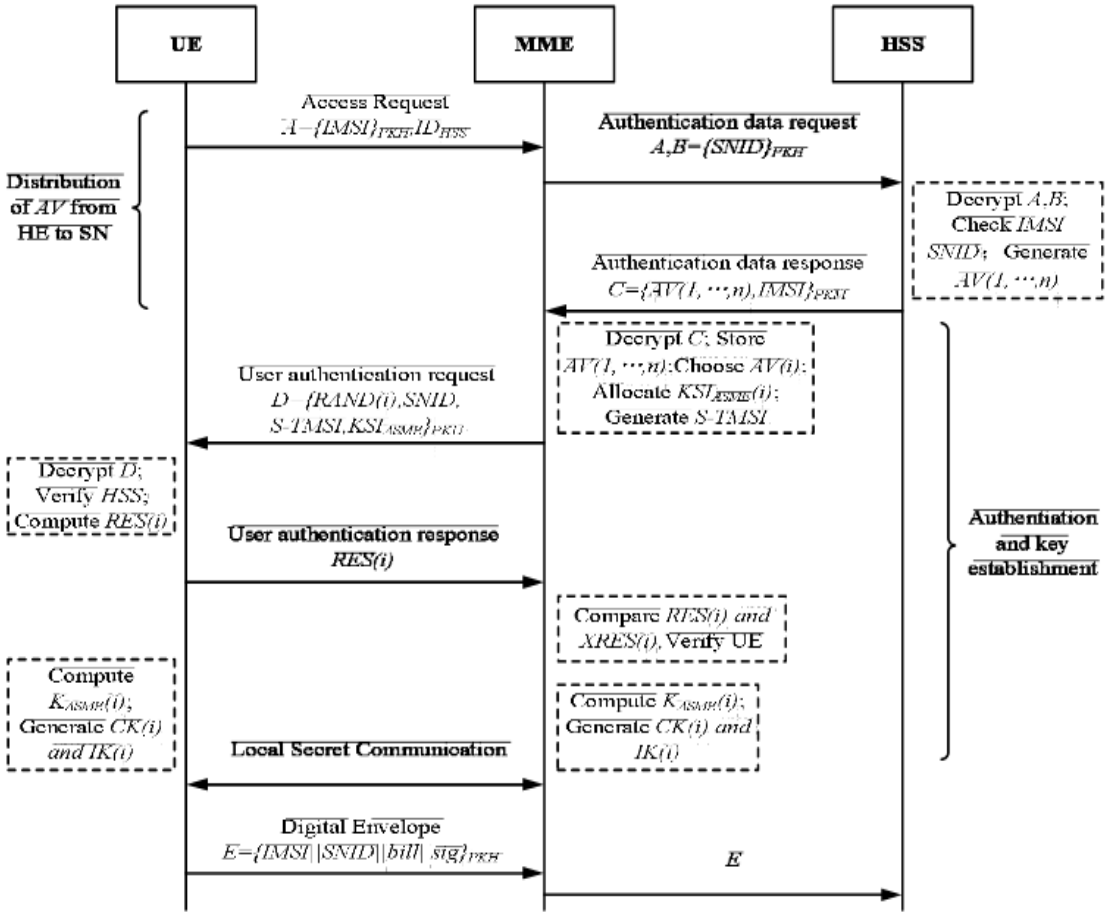

Figure 29. SE-EPS AKA process, from [27].

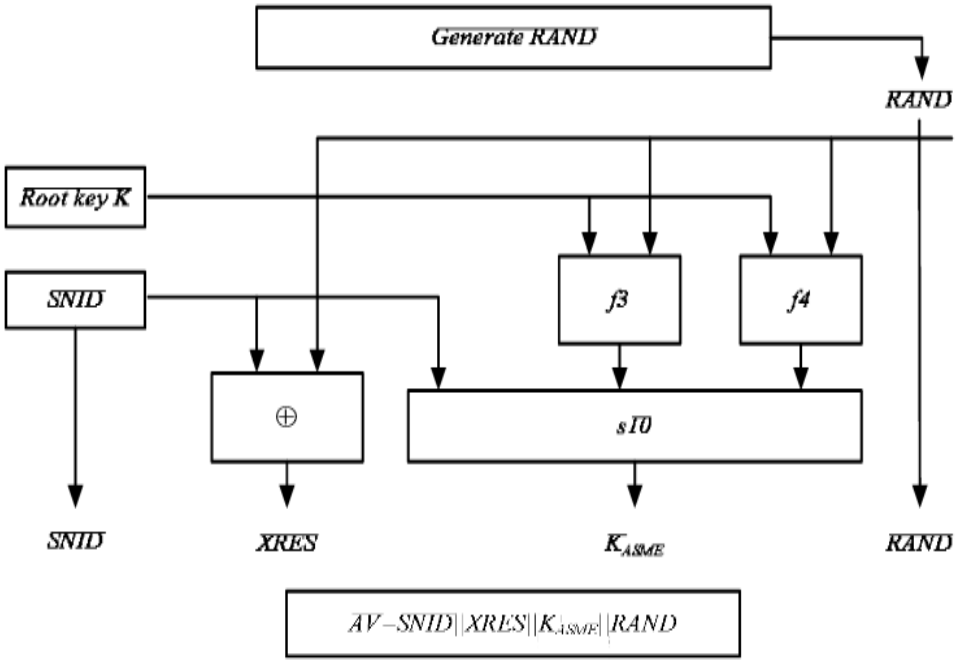

Figure 30. The SE-EPS AKA authentication vector generation algorithm, from [27].

The MME decrypts the message, derives the AV array and IMSI and stores the AV array in its database. Thereafter, it selects a non-previously used AV (i) and extracts the corresponding RAND (i) and the SNID. The MME allocates the cipher Key Set Identifier, KSI_{ASME} (i), to K_{ASME} (i) of the AV (i) and by using a shared algorithm between the MME and the UE, as well as the IMSI, generates the S-TMSI (SAE-TMSI). After one-time authentication and cipher key negotiation takes place between UE and MME, they both store the relationship between the KSI_{ASME} (i) and K_{ASME} (i) that can be used for future verification without the UE and the SN having to follow the initiating authentication process. The MME encrypts RAND (i), S-TMSI, KSI_{ASME} (i), and SNID with the public key of the UE and sends it as a user authentication request to the UE. After decrypting the received message with its private key, the UE derives the values from the received message, computes the S-TMSI, and authenticates the HSS by comparing the calculated and received S-TMSI. If the HSS-provided values

are verified, the UE computes the RES (i) and sends it to the MME as an authentication response. Otherwise, the process terminates. The MME verifies the UE by comparing the RES and XRES and, if valid, both the MME and the UE compute the K_{ASME} (i), consider it as an intermediate key cipher key, and compute the CK and the IK to provide confidentiality and integrity, respectively. If RES and XRES are not equal, the process is terminated. Finally the MME and the UE store the relationships between the S-TMSI and the arrays (IMSI, AV (1,2,…,n), CK(i), IK(i),KSI_{ASME} (i), K_{ASME} (i)) and (IMSI, SNID, CK(i), IK(i),KSI_{ASME} (i), K_{ASME} (i)). There is one more optional step in which the UE uses its private key to sign the IMSI, SNID, and billing information and sends it encrypted, with HSS public key, PK_H, to the HSS through the MME for future accounting information.

Thus, the security of the user's identity and the exchanged information is enhanced by the use of SE EPS-AKA [27].

b. EC-AKA II Protocol

The EC-AKA2 protocol [26] is an improvement of EC-AKA, and it is a new solution that came out in 2013. It offers a solution that mostly addresses IMSI catching and user tracking vulnerabilities. It provides confidentiality and integrity protection, and it uses symmetrical encryption, which is faster than asymmetrical encryption, thereby reducing the delay imposed by security mechanisms. The EC-AKA II procedure is depicted in Figure 31.

The procedure starts with the UE generating three random keys: RandomEncKey, RandomIntKey and RandomUESecCapab1. It concatenates these with the IMSI and UESecCapabilities. The result is encrypted with the key, TIK. This key is the result of the concatenation of PIK and RandomIntKey, where PIK can be generated by a hashing function from the IMSI and is the permanent, pre-shared, integrity key (PIK) between the HSS and the UE. The message containing these values and the SNID are encrypted with the public key of HSS,

PK_H, and the resulting message, called the NAS attach request, is sent along with the HSS identity, ID_{HSS}, to the MME.

The MME compares the received ID_{HSS} to check if it is in the list of the HSSs that implement EC-AKA. If not, it treats the request as an AKA to ensure backward compatibility. If it does find the HSS identifier in its list but the message is sent unencrypted, the process is terminated since the security requirements are violated. If the identifier is valid and the message is encrypted, the MME detaches the $ID_{HSS,}$ adds the SNID encrypted by the PK_H and sends it along with the received encrypted message to the HSS.

The HSS, after receiving the message from the MME, decrypts both the two encrypted messages and compares the SNIDs to verify that they are the same. The private key, which is extracted from IMSI, along with a random number, RAND, are used to generate the authentication vectors that are going to be used not only for the current process but also for future authentication processes. The TIK is generated in the same way as on the UE. The original encrypted message from the UE is checked for integrity and if it passes the test an encryption key, EK, is generated by applying the XOR logic function to the PEK and the RandomEncKey. The RandomEncKey was derived from the decryption of the message received by the HSS from the MME. If the integrity check fails, HSS rejects the request and sends back an error message. If it passes the integrity check, the procedure continues by the HSS applying an integrity check to the all the previously mentioned with the key, IK, concatenating the result with the result of integrity check and encrypting it with serving network's public key, PK_M. The encrypted result is sent back to the MME.

The MME, after receiving the authentication response from HSS, decrypts it and authenticates the network; otherwise, it terminates the procedure. The MME derives the AV (1) and it integrity checks the RAND (1), AUTN, and KSI_{ASME}. The integrity check result is concatenated with RAND (1), AUTN and $KSI_{ASME.}$ Then an algorithm from the ones specified in UESecCap with the key, EK, is chosen in order for the result noted earlier to be encrypted. Then a random

key, RandomUESecCapab1, is added to the chosen algorithm and it is XORed with the encrypted result of the concatenation. The result is sent as a user authentication response to the UE.

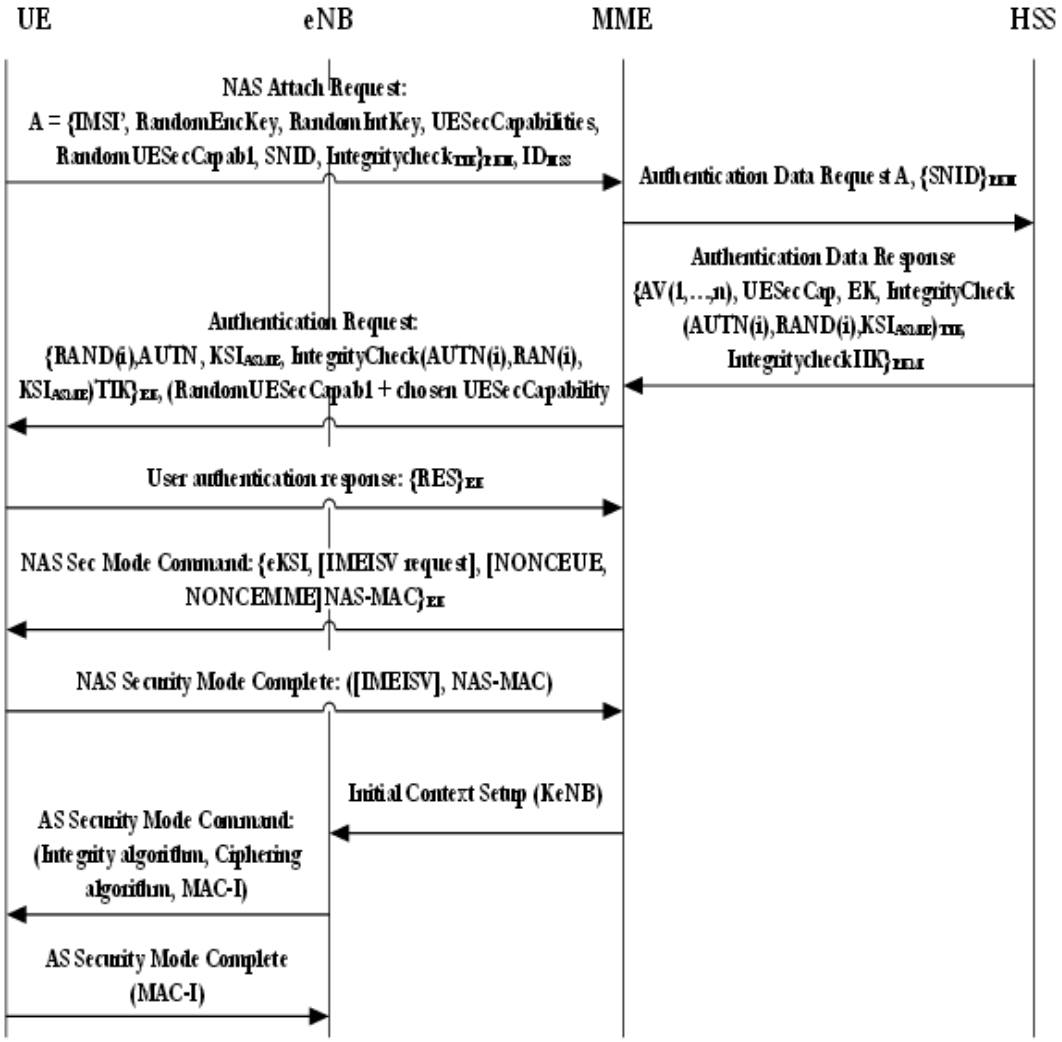

Figure 31. EC-AKA II procedure, from [26].

The UE, after receiving the message, extracts the encryption algorithm, decrypts the data, and by checking the RAND (1), AUTN and KSI_{ASME}, verifies the network's authenticity. If the verification fails, it terminates the process. Otherwise, it computes the RES and sends it, encrypted by the

temporary encryption key, EK, to the MME. The MME then compares the RES with the XRES and, if they are the same, authenticates the UE. Then the keys that are generated through the AKA are used to protect the communication integrity. The MME, using the same algorithm and temporary key EK, encrypts the message needed to establish the secure communications along with the integrity checking result and sends it to the UE using the NAS Secure Mode Command. The encryption and integrity algorithms are selected from the list of UE security capabilities that were sent from the UE. The UE verifies the receipt of the message by replying with a NAS Secure Mode Command complete message.

Finally, the MME connects to the eNB to which the UE is attached and sends all the necessary information, like keys and algorithms, needed to establish integrity and confidentiality protected AS communication. The message achieving this is called the Initial Context Setup message, and it does not need any protection since the communication between the MME and the eNB is sufficient using IPSec. Lastly, the eNB orders the UE to start using AS Security Mode command and the UE responds informing the eNB about the start of integrity and confidentiality protection enforcement.

The procedure is completed and all the messages from that point are protected, eliminating in this way many of the LTE weaknesses [26].

3. WiMAX Security Enhancement

In order to protect against the vulnerabilities of WiMAX security a few methods are introduced.

a. Management Messages Authentication Solution

By utilizing HMAC or CMAC digests, the non-authenticated messages that are identified in the WiMAX vulnerabilities analysis can be authenticated. Of course, there is a debate between security and performance since the authentication may need up to 168 bits. The more optimal solution is to

secure only those unauthenticated management messages causing serious defects if forged; the other messages can remain unauthenticated. In order to keep as minimum as possible the size of the messages, the Short HMAC or CMAC can be used. The Short HMAC provides a digest of 104 bits. The alternative solution, CMAC, is based on AES128, resulting in a 128-bit value. However, it can be truncated to 64 bits and finally ends up being 104 bits when all additional information is added [30].

Broadcast messages, however, use symmetric encryption and share the key among all member groups. This generates a potential for forged messages from any member of the group. However, using symmetric encryption outside the group may increase the security and the speed of processing, mitigating the risk but not offering complete protection [30].

Another solution is the utilization of asymmetric security. The private key of the BS is used to sign the management messages. The mobile stations verify the received messages by use of the associated public key. However, asymmetric encryption is much slower and requires a management mechanism for the asymmetric keys such as a certificate authority [30].

b. Unencrypted Management Communication Solution

Enforcing confidentiality and preventing attackers from reading management messages requires management messages be encrypted. After the authentication procedure is completed, the common key agreed upon by the authenticated parties may be used to enforce confidentiality. Thus, the TEK exchange and all messages that follow can be encrypted. However, either a security association for every management connection or a global management security association is required to eliminate Authorization Keys having to be frequently updated.

Encryption should be enforced as early as possible. A Diffie-Hellman (DH) key management protocol may be used to mitigate this vulnerability, a vector for Man-in-the-Middle attacks. Through this protocol the

initial management messages are protected. The DH protocol is a symmetric encryption that includes key exchanges between the MS and the BS.

The following equations express this protocol. For these, ‘a’ is the private key of the MS, ‘b’ is the private key of the BS, ‘P’ and ‘G’ global variables, and ‘G’ is a primitive root of ‘P’. Thus, the public keys of the MS and the BS are correspondingly:

PK_{MS}= G^a modP

PK_{BS} = G^b mod P

In the following diagram the four steps of the DH protocol are depicted:

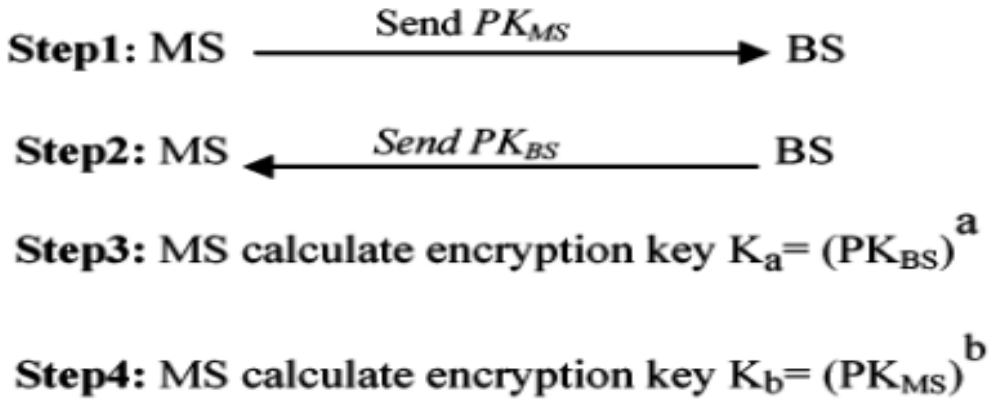

Figure 32. DH four step key exchange protocol, from [29].

It can be algebraically proven that K_a and K_b are equal. Thus, the encryption scheme used is symmetric. The encryption process that is used is depicted in the following diagram and it is consists of a bitwise XOR operation of plaintext and the key generated by DH.

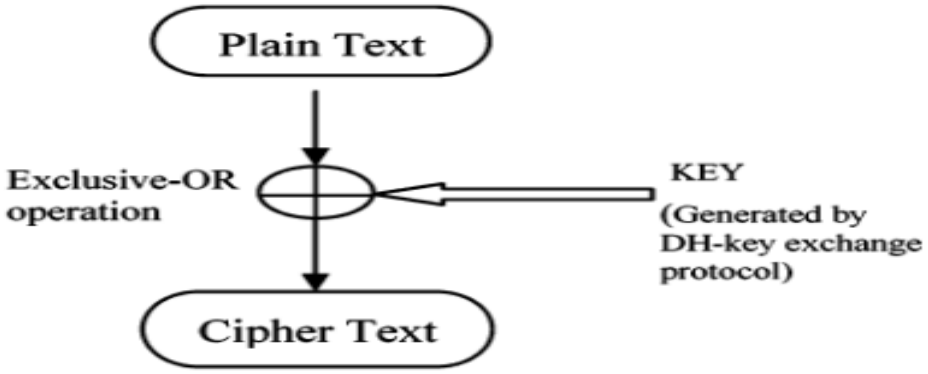

Figure 33. Encryption process by using the key generated by DH algorithm, from [29]

In this manner, the management messages may be encrypted even before the authentication process is completed [29].

c. Shared Keys in Multi- and Broadcast Service

If a shared key is used for the Multi and Broadcast Service, every member can potentially forge messages. Scientists proposed solutions to secure the update command message of key distribution so that attackers have no access to messages. Three distinct approaches are analyzed [30].

(1) Avoid broadcast key updates. This approach suggests using a unicast method of sending the updates to individual MS. Then the key is encrypted using the KEK which is available only to the distinct MS. Before the GTEK expires the BS sends the update command to each MS. This procedure is depicted in the left part of Figure 34 showing the similarity with GKEK update command message.

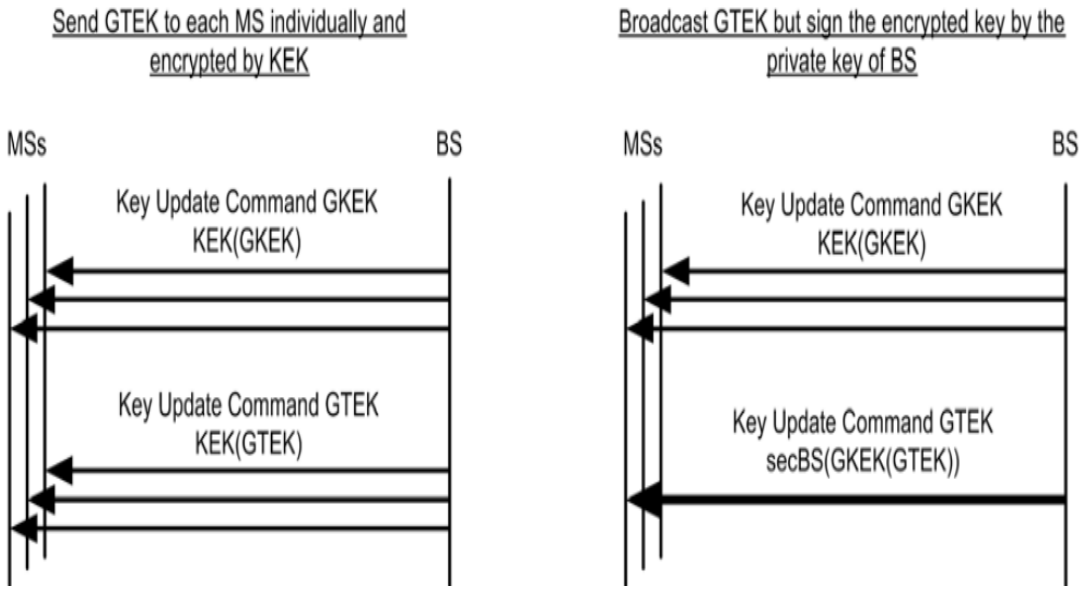

Figure 34. Potential solution for secure GTEK transmission, from [30].

(2) Public key cryptography. In this approach the BS broadcasts the GTEK update message, which is encrypted by the shared key GKEK and then signed by the private key of the BS. Then, each MS, after receiving the message, verifies the signature of the BS and then decrypts it by using the GKEK. This procedure is depicted in the right part of Figure 34.

(3) GTEK hash chain. In this approach the BS generates a random number, the $GTEK_0$. By applying a one-way hash function, f, to the previous GTEK value, each subsequent GTEK value is produced, as follows [29]:

$GTEK_0$ = random ()

$GTEK_1 = f(GTEK_0)$

$GTEK_2 = f(GTEK_1)$

$GTEK_n = f(GTEK_{n-1})$

The only key in this chain authentication that cannot be authenticated is the last one, the $GTEK_n$ that has to be sent securely to every

MS. One secure way is for the BS to send $GTEK_n$ through a unicast update message of the GKEK encrypted by the KEK.

After receiving the new GTEK, the MS uses the one-way hash function to verify its integrity. If the authentication is positive, the new GTEK overwrites the previous one. Otherwise, the message is discarded and a new GTEK is requested by the MS, as depicted in the following figure.

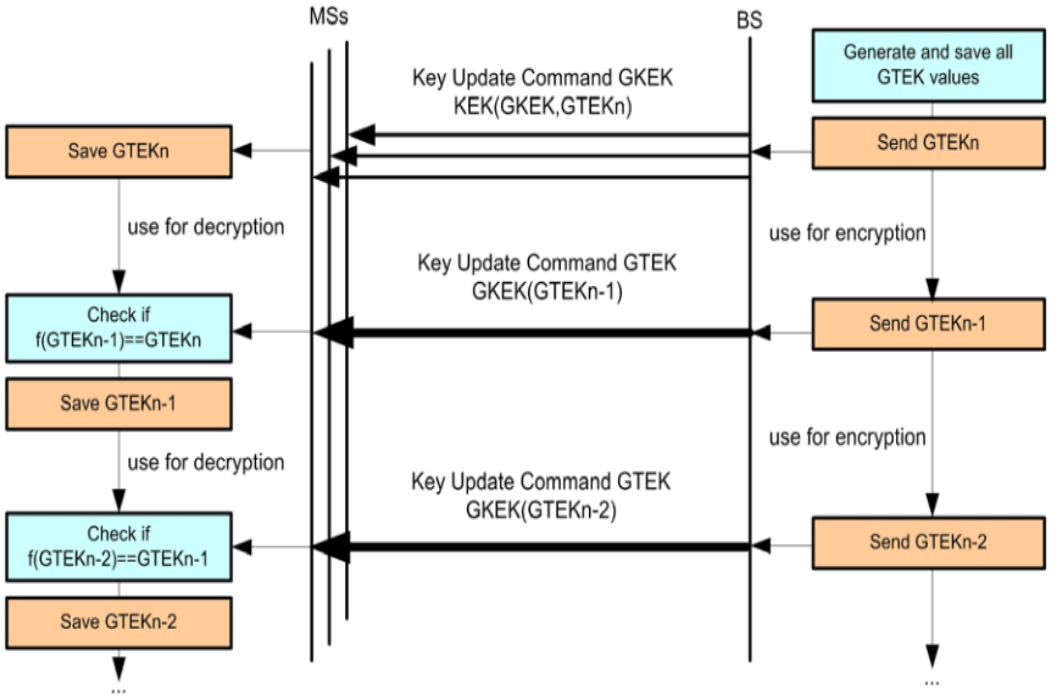

Figure 35. GTEK hash chain solution, from [30].

In the next chapter an evaluation of these solutions will be provided to understand how effective they may be for making communications more secure. Additionally the impact of security on military communications will be analyzed.

IV. ANALYSIS OF SUGGESTED SOLUTIONS AND SECURITY IMPACT ON MILITARY COMMUNICATIONS

Chapter III presented an analysis of security issues in the three WAN wireless technologies of interest to us and a few solutions to address their security concerns. In this chapter an evaluation of these solutions is provided. A discussion of how the security issues may affect military communications completes the chapter.

A. UMTS SECURITY SOLUTIONS EVALUATIONS

1. Defense Strategy Model Evaluation

The defense strategy model [21] aims to eliminate some of security access domain vulnerabilities. The two main security targets that the solution is concentrated on are the mutual authentication between the user and the network, and the establishment of the cipher and integrity keys after a successful authentication. In the proposed model, concern about the SN authentication is addressed. The introduction of the nonce, R1, enforces successful identification of the HN. The introduction of the nonce, R2, enforces successful identification of the SN. The successful authentications are also based on the fact that the messages exchanged among the three entities, the MS, the SN, and the HN, are encrypted. Specifically, the authentication between the MS and the HN is enforced by using the shared key, KHM; between the HN and the SN by using the public key of SN; between the SN and the MS by using the private key of the SN; and between the MS and the SN by using the public key of the SN. The asymmetric encryption schemes that are used, even if they are slower, offer better security. In this manner, the MS and the HN recognize the SN and make sure that the messages they exchange come through the legitimate SN and not a malevolent attacker. Thus, the Man-in-the-Middle attacks, along with the DoS attacks, are prevented in the network, enforcing confidentiality and integrity of the exchanged messages [21].

2. Enhanced Enhancement Mobile Security and User Confidentiality for UMTS (EMSUCU) Evaluation

The Enhanced EMSUCU [18] applies three measures beyond those taken by the EMSUCU solution to enforce additional security. The first is the introduction of a temporary key, TK, instead of using the secret key, K; the latter is only used to generate a new temporary key every time an EMSUCU is executed. Moreover, the key, TK, is transferred securely on the radio access since it is encrypted. The second measure is that of using CK and IK one last time before they expire to initiate the AKA procedure; this provides protection to the messages exchanged during the AKA procedure. The third is the establishing the size of 128 bits for CK, IK, and TK and the size of 256 bits for K, thereby making it more difficult for an adversary to attack. These three countermeasures prevent cipher-text and known-plaintext attacks that an adversary could leverage to retrieve the keys, as it is described in Chapter III. By protecting K, forward and backward security is enforced, since K's disclosure would have allowed the attacker the ability to decrypt past and future communication.

Furthermore, the Enhanced EMSUCU maintains the advantages of simple EMSUCU. The IMSI and the TMSI are sent encrypted, providing user confidentiality. These countermeasures prevent Man-in-the-Middle and rogue/fake BS station attacks [18].

3. Secure-Authentication Key Agreement Protocol

The S-AKA [20] provides protection from rediretion attacks. The Location Area Identity, which identifies the location of the BS, is sent in clear in the UMTS AKA. An attacker can exploit it and mount a redirection attack to forward traffic toward a fraudulent BS. The attacker can gain access to the messages then. In the S-AKA protocol the LAI is protected. Using a message authentication code, the S-AKA enforces integrity protection. Thus, the LAI of the BS can not be compromised and the HN can verify that the BS is a legitimate one. Preventing redirection attacks offers the user protection against being charged for services

through a network other than his home network. Moreover, it prevents the user from being compromised since it protects him from being connected to a network that has weak or no encryption.

Additionally, the S-AKA offers resistance to manin-the-middle attacks. The new encryction key, PLK, is introduced in case the MS is connected to a GSM BSS. The MS encrypts the payload with the PLK and provides communication confidentiality with the SGSN. Thus, there is no risk of messages being eavesdropped on or modified over the wireless network.

Even if it is not a security advantage, it should be noted that S-AKA reduces bandwidth consumption. The HLR/AuC, after authenticating the MS, sends to the SGSN a delegation key, DK, which enables the SGSN to authenticate the MS in future communications without having to interact with the HLR. In this way, the messages exchanged between the SGSN and the HLR are reduced and thereby the bandwidth consumption, as well [20].

B. LTE SECURITY SOLUTIONS EVALUATIONS

1. Security Enhanced Evolved Packet System–Authentication Key Agreement

In the Security Enhanced EPS-AKA (SE EPS-AKA) [27], the IMSI is encrypted using the public key of the HSS, PKH. The encryption provides security for the IMSI, making sure the IMSI can’t be disclosed to an adversary. Using encryption, the Man-in-the-Middle attacks that were analyzed in Chapter II are not possible any more. Moreover, via IMSI encryption protection, the subscriber location remains confidential.

The serving network identity, SNID, in SE EPS-AKA is encrypted using the PKH when it is transmitted thereby avoiding being disclosed to adversaries. Since it is encrypted, the attacker cannot derive the SNID and cannot mount a false BS attack or fraudulent network attack.

In the SE EPS-AKA, the SQN is not needed to maintain the freshness of the AV. Even if the attacker eavesdrops on the messages exchanged between

the MME and the HSS or the UE, he cannot decrypt the messages. Thus, the public encryption that is enforced in SE EPS-AKA provides more security and eliminates the need for the SQN functionality [27].

2. Ensured Confidentiality Authentication and Key Agreement II (EC-AKA2) Protocol

The Ensured Confidentiality Authentication and Key Agreement II protocol [26] offers confidentiality. The IMSI is not sent in the clear at all. It is encrypted, instead, with the public key of the HSS, making the IMSI's disclosure impossible. Moreover, when disclosure of the UE identification to MME is needed, the UE has to send his IMSI; however, EC-AKA2 suggests using the last TMSI instead of sending the IMSI or IMEI, protecting both values. This enforcement prevents attackers from compromising the IMSI or mounting certain DoS attacks.

When a new eNB has no access to an old eNB it requests the associating user‘s IMSI. EC-AKA2 adopts a new technique compared to EC-AKA and EPS-AKA. It forces the user to re-run the EC-AKA2 instead of sending the IMSI. This technique provides further IMSI protection. The attacker is not able to exploit the IMSI-TMSI relationship to track of the user movement. In this way, false BS attacks become infeasible and confidentiality is enhanced.

Another functionality that EC-AKA2 offers is the TMSI reassignment for all users in a cell. This functionality further prevents an attacker from linking an IMSI to its associated TMSI. Even if the attacker knows the previous MSI/TMSI relationship, after TMSI reassignment the attacker will have no idea about the new relationship. The period of reassignment is specifically defined by the protocol designer.

Last, the Service Network Identity, SNID, is protected in EC-AKA2. The SNID is encrypted in the first message sent from the UE to the MME. Then the MME sends the SNID, encrypted by PKH, and the encrypted message from the UE that also includes the SNID to the HSS. The HSS decrypts the received message and compares the SNIDs to verify the authentication of the SN. The

HSS sends back a message to the MME encrypted with the public key of the legitimate SN. Even if an attacker impersonates a legitimate SN, he does not have the SN's private key to decrypt the last message received from HSS, preventing authentication. This functionality is feasible with EC-AKA, too, and enhances secure authentication of the SN, preventing false base station attacks [26].

C. WIMAX SECURITY ENHANCEMENT

1. Management Messages Authentication Solution Evaluation

According to [30] the use of CMAC or Short HMAC provides security to management messages exchanged and reduces the size of the message to the minimum possible of 104 bits. In this way the effectiveness of the protocol is overly affected.

Using symmetric encryption, the broadcasted messages are protected and the encryption is still fast maintaining the protocol to a sufficiently effective level. Moreover, the protocol is secured from attackers other than group members.

The asymmetric encryption may be the last option if the designers consider the protocol performance essential. The key management mechanism needed along with the time it takes for the asymmetric encryption to be executed make it less than the most attractive solution.

All three solutions provide integrity for the protocol. The advantages are many, since there are a lot of cases in which an attacker could benefit by an unauthenticated message and interrupt the communication between the MS and the BS. By protecting the management messages, a potential attacker is prevented from generating an unauthenticated message, thereby waking up the MS to receive traffic. Thus, the attacker can no longer keep the MS active and stress its battery, nor can an attacker use the relevant message to control the transmitting power of the MS. In this way, the adversary is prevented from setting the transmitting power either to a minimal level such that the MS cannot be recognized by the BS, or to an excessive level to stress its battery. By

authenticating the message that advertises the neighboring BS, the attacker has no option of omitting or modifying information about such neighbors or even advertising a non-existing BS to an MS [30].

2. Unencrypted Management Communication Solution Evaluation

The Diffie-Hellman protocol [9, 29, 30] that is used to enforce confidentiality is a symmetric encryption scheme: that means it is fast compared to the asymmetric schemes. Second, it protects the management messages from being disclosed to an adversary. In this way, useful information that could be used to mount an attack is not made available to an attacker. For example, by eavesdropping on bandwidth-related management messages an attacker may surmise the importance of a particular user by recognizing how much bandwidth the user is allocated. By encrypting authorization requests, which include user security capabilities, the adversary would no longer have insight into which user needs more support or stronger protection. Another advantage of encryption is that it prevents disclosure of the user and connection QoS parameters. Since the QoS parameters include information about service and user priorities, an attacker could otherwise get insight into mission priorities if the messages were not encrypted. Last, message encryption during the initial ranging procedure protects the user from being tracked. Encrypting the messages during ranging leaves no loophole for the attacker to calculate the location of the user. Thus, the attacker is denied all the useful information he used to derive by eavesdropping on unencrypted messages [9, 29, 30].

3. Shared Keys in Multi- and Broadcast Service Solution Evaluation

The solutions provided in [28, 30] help to avoid attacks by addressing a vulnerability exposed when the BS sends a message to every member of the group using the same Group Key Encryption Key (GKEK). By using a unicast

message to update the GTEK, the message sent from the BS to a distinct MS can only be decrypted by the legitimate MS. Nobody else knows the KEK but the individual MS.

Further, by using public key cryptography, the BS signs the encrypted GKEK update message using the newly updated GTEK. In this way, the protection of a legitimate BS is enforced since the MS can verify that the BS that sent the message was the legitimate one.

Using the GTEK hash chain, all of the keys in chain authentication are authenticated except the last one. The last key, GTEKn, though, is protected since it is sent encrypted through the unicast GKEK update message.

Each of these solutions offers protection against a potential group member attacker that wants to fool the rest of the group members. The attacker cannot impersonate a BS. Moreover, the attacker is prevented from using the GTEK update command message to mislead the rest of the group members; the attacker cannot distribute a falsely modified GTEK to the other members any more. Thus, the risk of members accepting a GTEK modified by the attacker and being unable to decrypt traffic from the legitimate BS is eliminated [28, 30].

D. SECURITY SOLUTIONS' CONTRIBUTION TO MILITARY COMMUNICATIONS

In military communications, a significant effort is made to improve not only the quality of service but also the security provided. The communications security is vital in military communications.

The communications demand on the battlefield now is not just voice communication through radio, as it used to be a few years ago. Today, military forces desire various fidelity pictures, video, and command and control systems while on the move, and all the messages that support these functionalities must be exchanged in a secure manner. Secure communications can make the difference between a mission success and a mission failure, life and death.

Military communication must ensure that the data exchanged are protected with respect to confidentiality and integrity, that only authorized people have access to the resources and that the military entities have critical data available at any time.

One of the most important aspects of security in military communications is mutual authentication. It has to be ensured that there is no chance of a user registering to a rogue BS. During mutual authentication, both the BS and the SS prove that they know the shared secret. In the event of a rogue BS, all data traffic going through that BS might be compromised to a potential attacker.

Another important aspect of mission critical communication is encryption. Though there may be cases in commercial communications where disclosure of data has a limited effect on the user or the network, with military communications information disclosure may be exploitable from the enemy in future missions even if the information has no meaning at that time. For example, methods of operation may be revealed that might be exploited in future engagements.

Another aspect that may be harmful in military communications is the disclosure of authorization requests. The authorization requests include the subscriber's capabilities. The attacker may assume which subscribers handle sensitive data by checking those subscribers requiring strong protection capabilities.

The disclosure of management messages is another weakness that may pose a great impact in military communications. An eavesdropper can collect valuable information about subscriber's station capabilities by management messages sent in the clear. Moreover, from QoS parameters of subscribers and connections, an attacker may get an insight about service and user priorities that may reflect the mission priorities [33].

More pertinent to the three technologies upon which this thesis focuses, IMSI catching is one of the issues of concern. Captured IMSI information may make it possible for an attacker to mount many distinct DoS attacks. The DoS

attacks may consume bandwidth, modify initial security capabilities of the MEs, or authentication parameters preventing network and user authentication, or simply making a lot of authentication requests and exhausting the serving network management resources and the HSS's computational power.

Another important area of risk is user equipment tracking. The attacker can track the user's temporary or permanent ID and use it to backtrack along the user's entire trace. Moreover, the attacker can impersonate a serving network and have access to encrypted data. This weakness is really important for military troops because it can disclose their location to the enemy [26].

The fact that the SN is not authenticated in 3G networks can lead to Man-in-the-Middle attacks in both wireless and wired networks that are potentially harmful for military communications. DoS attacks may make the system unavailable to the user, adversely impacting command and control actions. This becomes harmful for troops that are in battlefield and suddenly lose communication.

The potential CK, IK and secret key, K, disclosure is another important weakness. An attacker may gain access to a significant amount of data if the CK and IK are disclosed. The worst scenario, though, is disclosure of the secret key that compromises the security of past and future communications. This could be devastating for operations.

Leveraging unencrypted and unauthenticated management messages, the attacker might harm communications. In the case of unencrypted messages, the adversary can eavesdrop on the critical information, derive useful information, impersonate an MS, or mount Man-in-the-Middle and DoS attacks. Then the attacker can relate the derived information with the user equipment, which is even worse in military communications. In the case of unauthenticated messages, the communication between the MS and the BS will be disclosed to the attacker, giving a big advantage to the attacker in military communications.

Each of these issues and examples highlights the criticality of the security of 3G, LTE and WiMAX, as many weaknesses may impact mission outcome. Thus, the solutions included in the second part of this chapter may become really beneficial for military communications.

After having analyzed some of the security weaknesses and attacks that these wireless technology face, a few solutions that may provide some degree of countermeasure were provided. The effect that these issues may have on military communications and the importance of security in those communications was depicted.

After having analyzed the vulnerabilities and solutions suggested in the literature as countermeasures to the deficiencies of the three cellular technologies, the evaluation of the methods and the impact on military communications was presented in this chapter. The final chapter summarizes the results and enumerates the security areas that were not addressed in this thesis, identifying areas for further study.

V. ANALYSIS OF SUGGESTED SOLUTIONS AND COMMERCIAL OFF THE SHELF (COTS) PRODUCTS

A. SUMMARY OF RESULTS

All the solutions provided, analyzed and evaluated in the previous chapters enforced security in wireless communications. Attacks such as Man-in-the-Middle, replay and Denial-of-Service were mitigated through these solutions. Through such solutions military communications can become more secure by protecting exchanged messages, subscriber identity and configuration capabilities and location information. Thus, the solutions provided can have a beneficial impact on military communications that can be critical for mission's success.

1. UMTS Security Solutions Performance

The solutions for UMTS security discussed in Chapter III provide methods to counter many vulnerabilities.

The first UMTS solution enforced mutual authentication between the user and the network to assure successful identification of the servicing and home networks. Moreover, it established cipher and integrity keys after successful authentication was completed. It also enforced encryption in communications between the mobile station and the network management entities. This solution offered security against Man-in-the-Middle and DoS attacks by enforcing confidentiality and integrity of the exchanged messages.

The Enhanced Enhancement Mobile Security and User Confidentiality for UMTS (Enhanced EMSUCU) protocol introduced a temporary key to protect the shared secret key when it is transferred on the radio access. The messages exchanged during initial steps of the AKA procedure were secured using the encryption and integrity keys to initiate the procedure before those keys expire. It also made the keys more secure by increasing the size of the keys used. In this way cipher-text and known-plaintext attacks were prevented. Furthermore, user

confidentiality was enforced by transferring the IMSI and TMSI encrypted thereby preventing man-in-the-middle and rogue/fake BS station attacks.

The Secure-Authentication Key Agreement protocol (S-AKA) offered protection against redirection attacks by enforcing integrity protection to the Location Area Identity of the BS. In this way, the legitimacy of the BS was established. Additionally, the protocol, by introducing a new key, enforced confidentiality to protect messages against eavesdropping or modification in the event a GSM BS is also involved in communications session. In this way, it prevented Man-in-the-Middle attacks over the GSM segments which normally provide no encryption.

2. LTE Security Solutions Performance

The Security Enhanced Evolved Packet System – Authentication Key Agreement (SE EPS-AKA) prevents man-in-the-middle attacks and subscriber location disclosure by encrypting the IMSI. Additionally, it prevents an attacker from mounting a false BS attack or fraudulent network attack by encrypting the Service Network Identity (SNID).

The Ensured Confidentiality Authentication and Key Agreement II (EC-AKA II) protocol prevents Man-in-the-Middle and DoS attacks by encrypting the IMSI in the first step of the authorization process and using a Temporary Mobile Subscriber Identity (TMSI) instead of the IMSI and IMEI in the remaining authentication steps. It prevents false base station attacks by encrypting the SNID. Additionally, it protected the IMSI/TMSI relationship and prevented false BS attacks by rerunning the EC-AKA II protocol whenever an eNB lost connection and had to reconnect and by periodically reassigning TMSIs for all users in a cell.

3. WiMAX Security Solutions Performance

The CMAC or Short HMAC algorithms were proposed to provide integrity since they authenticate the management messages. The symmetric encryption,

with the key shared among all member groups, or asymmetric encryption, using the private key of the BS, were two alternative ways of authenticating the management messages. Both solutions prevented data messages between an MS and the BS from being intercepted thereby inhibiting Man-in-the-Middle and DoS attacks.

The Diffie-Hellman symmetric encryption scheme was used to encrypt management messages. It prevented attackers from eavesdropping on messages that include user security capabilities, QoS parameters or bandwidth related information. In this way, it prevented MS impersonation, user tracking, and Man-in-the-Middle and DoS attacks.

Similarly, the unicast message method ensured that only the legitimate MS could decrypt the message containing the Group Key Encryption Key (GKEK) update. The public key cryptography ensured that an MS could verify the legitimacy of the BS that sent it the GKEK update message. The Group Traffic Encryption Key (GTEK) hash chain authentication also ensured that the GTEK keys used to update the GKEK message were generated and transferred securely. Each of these solutions would provide protection against a potential group member attacker trying to modify the GTEK and making the rest of the group members unable to decrypt traffic received from a legitimate BS.

B. FUTURE WORK

The thesis addressed issues having to do with authentication, authorization, and key distribution. Its focus was mostly restricted to security enforced through various authentication protocols and techniques to enforce the confidentiality and integrity of messages exchanged across the radio interface. However, other areas bear consideration; following are a few topics that should be explored in further studies.

Mobile Virtual Private Networks (MVPNs) are commonly used, especially in military communications. MVPNs are network configurations by which mobile users can have access to the home network resources as “locally-connected

users” even if they change location. MVPNs are based on the functionality of VPNs that use tunneling, authentication and encryption to provide a virtual local-access line to the hosting network over which entities communicate securely. MVPNs add benefit over fixed VPNs as they provide continuous service to the users even though the mobile user transitions across different technologies or connections. The logical IP address assigned by the MVPN service supports mobility as the user device may roam and switch across different technologies and networks while the service maintains the logical IP-address association to the actual IP-address, such as through the use of Mobile-IP, allocated to the host as it traverses different network segments. This offers desired flexibility to mobile users.

Another useful technique that offers benefit is multilayered security. The security is applied in layers by which security provisions or protocols are applied in tandem to analyze and enforce security. The security objectives, or dimensions, such as authentication, non-repudiation, confidentiality and integrity, are addressed to counter threats and potential attacks. One example of concerted efforts to provide for multilayered security is the ITU-T X.805 standard, which provides a multilayered, end-to-end, network-security framework across eight security dimensions to defeat threats.

Finally, efforts could be made to increase the security within the application layer. Adding intrusion detection systems and firewalls for gateways that track or control application traffic and enforce access control to specific applications will help to make the network less vulnerable to attacks. In summary, this thesis explored security issues pertinent to 3G, 4G/LTE and WiMAX networks; analyzed known vulnerabilities for each of these wireless technologies; assessed a few proposed solutions to make these technologies more secure and discussed the benefits of enforcing cellular security for military communications.

LIST OF REFERENCES

[1] C. Blanchard, “Security for the third generation (3G) mobile system,” Elsevier Science, Ltd., Philadelphia, PA, Information Security Technical Report, vol.5, no. 3, 2000.

[2] H. Imai, M. Rahman, and K. Kobara*, Wireless Communications Security*. Boston, MA: Artech House, 2006.

[3] N. Boudriga, *Security of Mobile Communications*. Boca Raton, FL: Auerbach Publications, 2009.

[4] 3rd Generation Partnership Project, 3GPP TS 35.206 V9.0.0 (2009-12), Specification of MILENAGE Algorithm set: An example algorithm set for the 3GPP authentication and key generation function f1*,f1,f2,f3,f4,f5,f5* [Online]. Available: http://cryptome.org/3gpp/35206-900.pdf

[5] M. Lei, H. Bi, and Z. Feng, "Security architecture and mechanism of third generation mobile communication," in *Proc. 2002 IEEE Region 10 Conference on Computers, Communications, Control and Power Eng.* (*TENCON '02)*, Oct. 2002, vol.2, pp. 813-816.

[6] 3rd Generation Partnership Project, 3GPP TS 35.201 V9.0.0 (2009-12), Specification of the 3GPP confidentiality and integrity algorithms, Document 1: f8 and f9 specification [Online]. Available: http://cryptome.org/3gpp/35201-900.pdf

[7] 3rd Generation Partnership Project, 3GPP TS 35.202 V9.0.0 (2009-12), Specification of the 3GPP confidentiality and integrity algorithms, Document 2: KASUMI specification [Online]. Available: http://cryptome.org/3gpp/35202-900.pdf

[8] R.C.N. Chiang, A. Sesmun, G. Foster, M. Young, and N. Baker. (May 2002). "Transport of mobile application part signaling over Internet protocol," *Communications Magazine, IEEE*, vol. 40, no. 5. pp. 124-128 [Online]. Available: http://ieeexplore.ieee.org.libproxy.nps.edu/stamp/stamp.jsp?tp=&arnumber=1000223

[9] N. Seddigh, B. Nandy, R. Makkar, and J-F Beaumont, "Security advances and challenges in 4G wireless networks," in *Eighth Annual International Conference on Privacy Security and Trust (PST),* 2010, pp. 17-19, 62-71.

[10] J. Cao, M. Ma, H. Li, and Y. Zhang, "A survey on security aspects for LTE and LTE-A networks," *Communications Surveys & Tutorials, IEEE*, pp. 1-20, April 2013.

[11] J. Feng, "Analysis, implementation and extensions of RADIUS protocol," in *International Conference on Networking and Digital Society (ICNDS '09)*, May 2009, vol.1, pp.154-157.

[12] Security in LTE and SAE-Network [Online]. Available: http://www.home.agilent.com/upload/cmc_upload/All/Security_in_the_LTE-SAE_Network.PDF?&cc=US&lc=eng

[13] Mobile WiMAX Part1: Overview and performance [Online]. Available: http://www.wimaxforum.org/news/downloads/Mobile_WiMAX_Part1_Overview_and_Performance.pdf

[14] J. Andrews, A. Ghosh and R. Muhamed, *Fundamentals of WiMAX: Understanding Broadband Wireless Networking*. Upper Saddle River, NJ: Prentice-Hall, 2007.

[15] P. Rengaraju, L. Chung-Horng, Q. Yi, and A. Srinivasan, "Analysis on mobile WiMAX security," in *2009 IEEE Toronto International Conference on Science and Technology for Humanity (TIC-STH)*, pp.439, 444.

[16] D. Caragata, S. El Assad, C. Shoniregun, and G. Akmayeva, "UMTS security: Enhancement of identification, authentication and key agreement protocols," in *2011 International Conference for Internet Technology and Secured Transactions (ICITST)*, Dec. 2010, pp. 278-282.

[17] M. Khan, A. Ahmed, and A.R. Cheema, "Vulnerabilities of UMTS access domain security architecture," in *Ninth ACIS International Conference on Software Engineering, Artificial Intelligence, Networking, and Parallel/Distributed Computing (SNPD '08),* Aug. 2008, pp. 350-355.

[18] D. Caragata, S.E Assad, I. Tutanescu, C.A. Shoniregun, and G. Akmayeva, "Security of mobile Internet access with UMTS/HSDPA/LTE," in *2011 World Congress on Internet Security (WorldCIS)*, pp. 272-276.

[19] M. Zhang and Y. Fang, "Security analysis and enhancements of 3GPP authentication and key agreement protocol," in *IEEE Transactions on Wireless Communications*, vol.4, no.2, pp. 734-742, March 2005.

[20] Y.L Huang, C.Y. Shen, and S.W. Shieh, "S-AKA: A provable and secure authentication key agreement protocol for UMTS networks," in *IEEE Transactions on Vehicular Technology*, vol.60, no.9, pp. 4509-4519, Nov. 2011.

[21] H. Li, S. Guo, K. Zheng, Z. Chen, Z. Zhang, and X. Du "Security Analysis and Defense Strategy on Access Domain in 3G," in *1st International Conference on Information Science and Engineering (ICISE)*, Dec. 2009, pp. 1851,1854.

[22] M. Al-Humaigani, D. Dunn, and D. Brown, “Security Transition Roadmap to 4G and Future Generations Wireless Networks,” in *Proc. 41st Southeastern Symposium on System Theory (SSST 2009)*, March 2009, pp. 94-97.

[23] M. Aiash, G. Mapp, A. Lasebae, and R. Phan, “Providing security in 4G systems: Unveiling the challenges,” in *Proc. Sixth Advanced International Conference on Telecommunications (AICT)*, May 2010, pp. 439-444.

[24] Y. Park and T. Park, “A survey of security threats on 4G networks,” in *Proc. IEEE Globecom Workshops*, pp. 1-6, Nov. 2007.

[25] U. Meyer and S. Wetzel, “A man-in-the-middle attack on UMTS”, in *Proc. of the 3rd ACM Workshop on Wireless Security*, Oct. 2004, pp. 90-97.

[26] J.B. Abdo, J. Demerjian, H. Chaouchi, and G. Pujolle, "EC-AKA2 a revolutionary AKA protocol," in *2013 International Conf. on Computer Applications Technology (ICCAT)*, Jan. 2013, pp. 1-6.

[27] X. Li and Y. Wang, "Security enhanced authentication and key agreement protocol for LTE/SAE network," in *7th International Conf. on Wireless Communications, Networking and Mobile Computing (WiCOM),* Sep. 2011, pp.1-4.

[28] C. Kolias, G. Kambourakis, and S. Gritzalis, "Attacks and countermeasures on 802.16: Analysis and assessment," in *IEEE Communications Surveys & Tutorials*, vol.15, no.1, pp. 487-514, First Quarter 2013.

[29] M.S. Rahman and M.M.S Kowsar, “WiMAX security analysis and enhancement," in *12th International Conference on Computers and Information Technology, (ICCIT '09)*, Dec. 2009, pp. 679-684.

[30] A. Deininger, S. Kiyomoto, J. Kurihara, and T. Tanaka, “Security vulnerabilities and solutions in mobile WiMAX,” *International Journal of Computer Science and Network (IJCSNS)*, vol.7, no.11, pp. 7-15, 2007.

[31] J. Al-Saraireh, S. Yousef, and M. Al Nabhan,”Enhancement of mobile security and user confidentiality for UMTS,” in *Second European Conf. on Mobile Government*, Brighton, Great Britain, 2006.

[32] S. Farrell, “The WAP Forum’s wireless public key infrastructure,” Elsevier Sciences, Ltd., Philadelphia, PA, Information Security Technical Report, vol. 5, no. 3, pp. 23-31, 2000.

[33] B. Bennett, P. Hemmings, and B.A. Hamilton, "Operational Considerations of Deploying Wimax Technology as a Last-Mile Tactical Communication System," in *2006 IEEE Military Communications Conference (MILCOM 2006)*, pp.1-7, 23-25 Oct. 2006

INITIAL DISTRIBUTION LIST

1. Defense Technical Information Center
   Ft. Belvoir, Virginia

2. Dudley Knox Library
   Naval Postgraduate School
   Monterey, California